VOLUME XIV—XV 1981 - 1982

UNIVERSITY OF THE PUNJAB

JOURNAL OF MATHEMATICS

(1882–1982) SPECIAL CENTENARY ISSUE

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF THE PUNJAB
LAHORE

Notice to Contributors

1. The Journal is meant for publication of research papers and review articles covering state of the art in a particular area of mathematical sciences.

2. Manuscripts should be typewritten and in a form suitable for publication. As far as possible, the use of complicated notations should be avoided. Figures, drawn on separate sheets of white paper in Black Ink, should be of a size suitable for inclusion in the Journal.

3. Contributions and other correspondence should be addressed to Dr. A Majeed, Mathematics Department, Punjab University, New Campus, Lahore-20, Pakistan.

4. The decision to accept or reject a paper for publication in the Journal rests fully with the Editoriai Board.

5. Authors, whose papers will be published in the Journal, will be supplied 30 free reprints of their papers and a copy of the issue containing their contributions.

6. The Journal which is published annually will be supplied free of cost in exchange with other Journals of Mathematics.

Printed by Haji Manzur Ahmad at Panchnad Industries & Press
1-Jan Muhammad Road, Anarkali LAHORE.

Published by Dr. M.H. Kazi for the University of the Punjab,
Lahore - Pakistan.

DR. MANZUR HUSSAIN (1921-1981)

Dr. Manzur Hussain, former Professor of Pure Mathematics and Head of the Department at Punjab University Lahore died on September 22, 1981 after a prolonged illness.

Dr. Manzur Hussain was the chief architect of the department of Mathematics at Punjab University. Being the first full-time Head of the department and one of the first few full-time teachers, he started from scratch in mid-fifties and through dedication, sense of purpose and hardwork built it into a full-fledged teaching department within a decade or so. He was deeply committed to the cause of mathematical education and impired his colleagues with the same zeal. He took pains in attending to his teaching work and administrative duties and through his personal example inculcated an abiding sense of duty and responsibility in his staff. He was always kind and considerate towards his students and attached a great deal of importance to their character-building through personal example and precept. He was equally kind and considerate towards his colleagues. He was very particular about the selection of new faculty members and merit was the sole criterion which guided his choice.

Here is a summary of Dr. Manzur Hussain's life and work. Born on September 15, 1921 in a village in the Jhelum district, he passed his matriculation and intermediate examinations in 1936 and 1938 respectively. He then obtained his B.A. with Honours in mathematics in 1940 and M.A. (Mathematics) from Islamia College Peshawar in 1942. He served as a lecturer in mathematics at Government College Sargodha and Government College Lahore till 1952 when he proceeded to Durham University, England for his doctoral work in Theory of Numbers. He studied at Durham University for about two and a half years (1952-54) and returned to Government College Lahore to serve again as a lecturer in mathematics. He joined the Mathematics Department of Punjab University as a reader on

June 4, 1956 and became its Head in February 1957. He was appointed Professor of Pure Mathematics on January 12, 1963. He served as Head of the department till the end of 1972. He was the author of several research papers and text-books. In a letter to him, Professor F.J. Dyson, a well-known mathematician made the following remarks about his research work, "your results are of striking beauty and elegance and comparable with the best of Ramanujan's work".

Dr. Manzur Hussain had a severe attack of paralysis on November 11, 1974. In a few days he recovered from the severity of the attack but it left him in a demented state. He made some improvement but never fully recovered. He remained on medical/privilege leave and was not allowed to join the department. His family suffered a great deal of hardship. The attitude of university authorities towards this Professor, who devoted the best part of his life to the service of University, left much to be desired. He retired from university service on May 24, 1979. In 1981 he had another severe attack of paralysis which proved fatal. He breathed his last on September 22, 1981.

May his soul rest in peace ! (Amin).

(Khalid Latif Mir)

DR. MUNAWWAR HUSSAIN (1936-82)

Dr. Munawwar Hussain, Head Mathematics Department Government College Lahore died in a road accident on the night of October 11, 1982. In his sudden and tragic death, the mathematical community of Pakistan has lost a decidicated and conscientious teacher, and a good research worker.

Born in 1936, Dr. Munawar Hussain received his school and early college education in his native town Sargodha. He received his B.Sc. with Honours in Mathematics degree from Government College Faisalabad in 1955 and then moved to Lahore and obtained his Master's degree in mathematics from the Punjab University in 1957, with distinction. Some of his teachers at this stage of his educational life were Dr. L.M. Chawala and Ch. Sultan Bux (Govt. College), Prof. Nasir-ud-Din and Prof. Sana Ullah Bhatti (Islamia College), Prof. F.D. Anjum Roomani (Dyal Singh College) and Prof. Maqbool Ilahi (F.C. College). He joined Government College, Satellite Town, Rawalpindi in October 1958 and served there till 1964, when he was transferred to Government College Sargodha. Munnwar Hussain was very keen to pursue research work in some field of applied mathematics and his desire did materialize in 1967 when he joined the first batch of M. Phil, students of the Quaid-e-Azam, University, Islamabad. Later he was registered for Ph. D. under the supervision of Prof. Q.K. Ghauri, who proved a source of great inspiration for his talented student. After completing his Ph. D. in 1971, Dr. Hussain joined the newly established Government College of Science, Lahore as an Assistant Professor. For a period of one year (1973-74) he served the Quaid-i-Azam University Islamabad before he joined Government College Lahore as Professor of Mathematics and Head of the Department.

The last nine years of Dr. Munawar Hussain's life were spent at Government College Lahore. His contribution in building the Mathematics Department of that college into a leading department will be long remembered. He inspired his colleagues through his personal example of hardwork, dedication and efficiency. Alongside teaching and administrative work, he carried on research in his field of specialization and wrote eighteen papers in the period 1973-1982. He also published a scholarly monograph entitled "Lagrange Equations of Motion". He was a reviewer of Mathematical Reviews and of journals published by Springer-Verlag. He played an important role in establishing the Punjab Mathematical Society and was elected its Vice-President in 1977. A few days before his death, he was busy in connection with the formation of All Pakistan Mathematical Association. He was an Associate Editor of the Journal of Natural Sciences and Mathematics published by Government College Lahore, and a Research Associate of International Centre for Theoretical Physics, Trieste.

May his soul rest in peace ! (Amin.)

[Muhammad Amin,
Mathematics Department, Govt. College Lahore.

and

Khalid L. Mir,
Mathematics Department, Punjab University, Lahore.]

The Punjab University
Journal of Mathematics
Vol. XIV—XV (1981-82), pp. 1-14

TOPOLOGICAL ALGEBRAS ISOMORPHIC WITH l_1 OR H (D)

By

TAQDIR HUSAIN*

Mathematics Department,
McMaster University, Hamilton, Canada and IAS, Australian National University, Canberra

Abstract

We give some necessary and/or sufficient conditions on bases of Banach or Fr'echet algebras with orthogonal or cyclic bases which make them isomorphic and homeomorphic with the standard algebras l_1 or H (D).

1. Introduction

Let E be a Hausdorff topological vector space (abbreviated to TVS) over the complex field C (or real field **R**, if needed). A countable subset $\{x_i\}$ of E is said to be a basis of E if for each $x \varepsilon$ E there exists a unique sequence $\{\lambda_i\}$ of scalars in **C** such that

1.1 $$x = \lim_{n\to\infty} \sum_{i=1}^{n} \lambda_i x_i = \sum_{i=1}^{\infty} \lambda_i x_i .$$

(See [2], [8], [11].)

If E, in particular, is a normed space, then 1.1 can be expressed as follows :

$$\left|\left| x - \Sigma_{i=1}^{n} \lambda_i x_i \right|\right| \to 0 \text{ as } n \to \infty .$$

(*) This paper was written when the author was visiting the Department of Mathematics, IAS, Australian National University, Canberra, Australia. This work was partially supported by a NSERC grant.

By the uniqueness of $\{ \lambda_i \}$, it follows that no member of a basis can be zero. Thus if $\{ x_i \}$ is a basis of E, then $x_i \neq 0$ for all $i \geq 1$. Moreover, by the dependence of λ_i's on x, it is clear that each $\lambda_i (x) = \lambda_i$ is a linear functional (called coordinate functional) for $i \geq 1$. In general, each λ_i need not be continuous. Whenever it is so, $\{ x_i \}$ is called a Schauder basis. By an application of the classical Open Mapping or its sister Closed Graph theorem, it can be shown that each basis in a complete metrizable locally convex TVS (in particular, for Banach spaces) is a Schauder basis. Schauder bases in Banach spaces (more generally in TVSs) are not unique. In other words, a Banach space may have more than one Schauder basis, if exist. As for existence, each separable Hilbert space does have a Schauder basis. There was a longstanding open problem in Functional Analysis whether or not each separable Banach space has a (Schauder) basis, but this was settled in the negative by Enflo [3] in 1973. There is a vast literature on bases in Banach spaces, for example, see [11] or [2]. The basis theory in more general topological vector spaces is also extensive. For example, see [8].

When the basis theory in Banach spaces and its applications in approximations were growing, there was an upsurge in the study of topological algebras, particularly Banach algebras. But there was no effort made to study the basis theory in topological algebras until this author and his coworkers looked at it. They showed that the basis theory in topological algebras could be equally interesting, if not more. For instance, see [4], [5], [6].

Let A be Hausdorff topological algebra over C, which means that an algebra A over C with a Hausdorff topology is a topological algebra if A is a TVS and the multiplication : $(x, y) \to xy$ is continuous. Note that some authors define topological algebras as those in which multiplication is continuous separately in each variable. But this distinction will not engage our attention here, since mostly in the sequal we will be dealing with complete metrizable topological algebras on which this distinction does not exist due to Aren's theorem (see [12] or [7]).

A countable subset $\{x_i\}_{i \geq 1}$ of a Hausdorff topological algebra A is said to be an orthogonal basis (*cf.* [4] or [5]) of A if $\{x_i\}$ is a basis of A as a TVS and $x_i x_j = \delta_{ij} x_i$ for $i, j \geq 1$ in which δ_{ij} is the Kronecker's delta. Clearly, $x_i x_j = 0$ if $i \neq j$ and $x_i^2 = x_i$ for all $i \geq 1$.

If one is interested in a weaker assumption of orthogonality of a basis $\{x_i\}$, one can assume only that $x_i x_j = 0$ for $i \neq j$, or that $x_i x_j = 0$ for $i \neq j$ and $x_i^2 \neq 0$, or that $x_i x_j = 0$ for $i \neq j$ and $x_i^2 = c_i x_i$ with some suitable conditions on c_i. For details, see [5].

Some of the consequences of our definition of orthogonality are : that a topological algebra A with an orthogonal basis is automatically commutative. For, if $x = \Sigma \lambda_i x_i$ and $y = \Sigma \mu_i x_i$, then $xy = yx = \Sigma (\lambda_i \mu_i) x_i$. Each coordinate functional λ_i is multiplicative : $\lambda_i (xy) = \lambda_i (x) \lambda_i (y)$ for $x, y \varepsilon$ A. Usually, a basis in a TVS becomes a Schauder basis only when the conditions for application of the Open Mapping and/or Closed Graph theorems are met, but an orthogonal basis in a topological algebra becomes a Schauder orthogonal basis under much less restrictions. Specifically, if A is a locally multiplicatively-convex (LMC, for short) algebra [9] *i.e.*, the topology of A is defined by a family $\{P_\alpha\}_{\alpha \varepsilon \Gamma}$ of seminorms satisfying the submultiplicativity condition :

$$P_\alpha (xy) \leq P_\alpha (x) \; P_\alpha (y)$$

for all $x, y \varepsilon$ A and $\alpha \varepsilon \Gamma$ and $\{x_i\}$ an orthogonal basis of A, then each λ_i is continuous [5]. In other words, $\{x_i\}$ becomes a Schauder basis of A. By using a delicate argument on the cardinality of the index set Γ, it is possible to show that each LMC algebra with identity and having an orthogonal basis is metrizable [5]. Hence, if A, in addition, is complete, then it is a Frechet (complete metrizable LMC) algebra with identity and having an orthogonal basis.

It is interesting to mention that (Schauder) orthogonal bases in topological algebras, unlike (Schauder) bases in topological vector spaces are essentially unique. That is, if $\{x_i\}$ and $\{y_i\}$ are two orthogonal (Schauder) bases of a topological (in particular, Banach) algebra A, then $\{x_i\} = \{y_i\}$, [5]. This general phenomenon is the extension of a well-known fact regarding the uniqueness of the orthogonal basis $\{e_i\}_{i=1}^{n}$ in $\mathbf{R}^n$ where $e_i = (0, \dots, 1, 0, \dots, 0)$ in which 1 occurs only on the i^{th} coordinate, otherwise each coordinate is 0.

It is also worth noting that the corresponding Orthogonal Basis Problem in a Banach algebra has a negative answer, because no non-commutative Banach algebra can have an orthogonal basis, since, as pointed out above, whenever a topological algebra has an orthogonal basis then it must be commutative. Thus, B (H), the algebra of all bounded linear operators on an infinite-dimensional Hilbert space H cannot have an orthogonal basis. Even for a commutative one ([3] and [5]), the answer is in the negative.

It is not only that an infinite-dimensional non-commutative normed algebra fails to have an orthogonal basis, but it is equally curious to note that an infinite-dimensional, commutative or not, normed algebra with an orthogonal basis fails to have identity. For, if $\{x_i\}$ is an orthogonal basis of an infinite-dimensional normed algebra A and if e is its identity, then $e = \Sigma_{i=1}^{\infty} x_i$. Since the series is convergent, there is a positive integer i_o such that for all $i \geq i_o$, $\| x_i \| < 1$. By orthogonality of $\{x_i\}$, we have $x_i = x_i^k$ and hence $\| x_i \| = \| x_i^k \| \leq \| x_i \|^k \to 0$ as $k \to \infty$ for all $i \geq i_o$ Hence $x_i = 0$ for $i \geq i_o$, contradicting the infinite-dimensionality of A. Hence A cannot have identity. This fact is also verified by the classical Banach algebras c_o, l_p $(1 \leq p < \infty)$ which have an orthogonal basis $\{e_i\}_{i \geq 1}$, where $e_i = \{\delta_{ij}\}_{j \geq 1}$ but no identity. Although an inifinite-dimensional normed algebra with an orthogonal basis cannot have identity, the same is not true for Fr'echet algebras.

Actually, most classical Frechet algebras with orthogonal bases do have identity. Incidentally, we have :

1.2 Theorem. ([4] and [5]).

Let A be a complete LMC algebra with an orthogonal basis. Then the following statements are equivalent :

(*i*) A has identity.

(*ii*) A is isomorphically homeomorphic with the Frechet algebra s of all complex sequences.

(*iii*) The map $\phi : x \to \{ \lambda_i (x) \}_{i \geq 1}$ of A into s is surjective.

Thus, up to isomorphism, there is only one Frechet algebra with an orthogonal basis and having an identity, and that is s. But this is not the case for B_o – algebras [12] (*i.e.*, complete metrizable locally convex algebras) with an orthogonal basis and having an identity.

1.3 **Example**

Let D be the open unit disk of the complex plane and H (D), the topological vector space of all holomorphic functions on D, endowed with the compact-open topology. If we define the product $f * g$ of f, $g \in$ H (D) by :

$$f * g\,(x) = \int_{|z| = r} f(z)\, g\,(xz^{-1})\, z^{-1}\, dz\,,$$

$|x| < r < 1$ (here $i = \sqrt{-1}$), then it is not difficult to verify that H (D) is a B_o – algebra with identity 1. Moreover, the sequence $\{1, z, z^2, \ldots)$ forms a basis of H (D) by the Taylor's theorem. It is rather routine to verify that $z^m * z^n = \delta_{mn}\, z^n$, *i.e.*, the basis $\{ z^n \}_{n \geq o}$ is an orthogonal basis. It is obvious that H (D) cannot be an LMC algebra and hence not a Frechet algebra, because otherwise it would be isomorphic to s which it is not.

Another feature of Frechet algebras with orthogonal unconditional basis is that they are functionally continuous [9], thus answering in the affirmative an old problem of automatic continuity of multiplicative linear functionals on a Frechet algebra with an unconditional orthogonal basis [4]. The problem for general Frechet algebras, however, still remains. See [7] for details.

2. l_1 – algebra

It is indeed well-known that the set l_1 of all complex sequences $a = \{a_i\}$ with $\Sigma_{i=1}^{\infty} |a_i| < \infty$ is a Banach space with pointwise addition, scalar multiplication and the l_1 – norm :

$$\| a \|_1 = \Sigma_{i=1}^{\infty} | a_i |$$

Actually, with pointwise multiplication, l_1 is even a Banach algebra without identity. The following facts about l_1 are known :

2.1 (*a*) l_1 is a commutative Banach algebra which has an orthogonal basis, $\{e_i\}$.

(*b*) $\{e_i\}$ is an unconditional basis of l_1 .

(Note : A basis $\{x_i\}$ in a TVS E is said to be unconditional if the series in each expansion $x = \Sigma \lambda_i (x) x_i$ ($x \varepsilon$ E) is unconditionally convergent [2]. It is equivalent to bounded-multiplier convergent [2].).

(*c*) The topological dual l_∞ of l_1 contains an element e such that $e(e_i) = 1$ for all $i \geq 1$.

(Note : $e = \{1\} \varepsilon l_\infty \setminus l_1$ and, by duality, if $a = \{a_i\} \varepsilon l_1$, and $b = \{b_i\} \varepsilon l_\infty$, then $< a, b > = b(a) = \Sigma_{i=1}^{\infty} a_i b_i$ is the value of the functional b at a.).

(*d*) The basis $\{e_i\}$ of l_1 is boundedly complete.

(Note : A basis $\{x_i\}$ in a normed space E is said to be boundedly complete if for all scalar sequences $\{a_i\}$, the boundedness of the sequence $\{ \| \Sigma_{i=1}^{n} a_i x_i \| \}_{n \geq 1}$ implies the convergence of $\Sigma_{n=1}^{\infty} a_i x_i$ to some $x \varepsilon$ E).

Now we have a characterization of those Banach algebras which are isomorphic to l_1 .

2.2 **Theorem**

Let A be a Banach algebra with an orthogonal unconditional basis $\{ x_i \}$. Suppose there exists $e \varepsilon A'$ (topological dual of A) with $e(x_i) = 1$ for all $i \geq 1$. Then A is isomorphic and homeomorphic with the Banach algebra l_1 iff the basis $\{x_i\}$ is boundedly complete.

Proof :

Indeed, if A is isomorphic and homeomorphic with l_1 , then, by identification, A will have two orthogonal bases $\{ x_i \}$ and $\{ e_i \}$ and therefore $\{ x_i \} = \{ e_i \}$, [5]. Since $\{ e_i \}$ is boundedly complete, the "only if" part follows.

To prove the "if" part, assume $\{ x_i \}$ is boundedly complete. Since A is normed, without loss of generality, we may assume that $\| x_i \| = 1$ for all $i \geq 1$. For each $x \varepsilon A$ there is a unique sequence $\{ \lambda_i (x) \}_{i \geq 1}$ such that

$$x = \Sigma_{i=1}^{\infty} \lambda_i (x) x_i .$$

Clearly, $\phi : x \rightarrow \{\lambda_i (x) \}_{i \geq 1}$ is an algebra homomorphism of A into s. Since the basis $\{x_i\}$ is unconditional, it is easy to verify that the series :

$$\Sigma_{i=1}^{\infty} \lambda_i (x) (sgn\, \lambda_i) x_i = \Sigma_{i=1}^{\infty} | \lambda_i (x) | x_i$$

is convergent in the norm of A and hence for each $f \varepsilon A'$ and in particular for $e \varepsilon A'$, we obtain the convergence of

$$\Sigma_{i=1}^{\infty} | \lambda_i (x) | .$$

This proves that $\{\lambda_i (x) \}_{i \geq 1} \varepsilon l_1$ *i.e.*, ϕ maps A into l_1 . By the uniqueness of scalars $\{\lambda_i (x) \}$ for each x, we see that ϕ is a one-one algebra homomorphism of A into l_1 . To prove that ϕ is onto, let $a = \{a_i\} \varepsilon l_1$. Since for all $n \geq 1$,

$$\| \Sigma_{i=1}^{n} a_i x_i \| \leq \Sigma_{i=1}^{n} | a_i | \| x_i \| \leq \| a \|_1 < \infty,$$

by the bounded completeness of the basis, there is $x \varepsilon A$ such that $x = \Sigma_{i=1}^{\infty} a_i x_i$. Hence $\phi(x) = \{ a_i \}$. Thus ϕ is an isomorphism of the normed algebra A onto l_1.

(Note : We have not used the completeness of A so far. To prove the homeomorphism of A onto l_1, completeness of A is used). Clearly, for each $x \varepsilon A$, $x = \Sigma_{i=1}^{\infty} \lambda_i(x) x_i$ implies $\| x \| \leq \Sigma_{i=1}^{\infty} | \lambda_i(x) | = \| \phi(x) \|_1$. Since ϕ is invertible, we see that $\| \phi^{-1}(y) \| \leq \| y \|_1$ for all $y \varepsilon l_1$. This proves the continuity of $\phi^{-1} : l_1 \to A$. By the Open Mapping theorem, we obtain the desired homeomorphism of A onto l_1. This completes the proof.

Note that the above theorem says that a normed algebra A with an orthogonal unconditional boundedly complete basis $\{x_i\}$ such that there is $e \varepsilon A'$ with $e(x_i) = 1$ for all $i \geq 1$ is isomorphic to the algebra l_1 and if A, in addition, is complete then it is also homeomorphic with l_1.

The Banach algebra c_0, the space of sequences converging to zero, also has an orthogonal basis $\{ e_i \}$. Since the dual of c_0 is l_1 and the basis $\{e_i\}$ of c_0 is shrinking [2], it is not difficult to find a similar characterization for c_0, using the shrinking property of a basis. Furthermore, since l_p $(1 < p < \infty)$ is a reflexive Banach algebra with an orthogonal basis $\{e_i\}$, it will be worthwhile to prove a similar characterization for l_p $(1 < p < \infty)$ using James theorem [2]. Some of these results will appear elsewhere.

3. H (D) – algebra

As mentioned above, the algebra H (D) of all holomorphic functions on the open unit disk D of the complex plane, endowed with the pointwise algebraic operations and compact-open topology, is a commutative Frechet algebra with identity. Actually, if we take D_r to be the open disk of radius $r > 0$ and centre 0, then H (D_r) is a commutative Frechet algebra with identity. In this section we want to show that an abstract commutative Frechet algebra with identity is isomorphically homeomorphic with H (Ω) for some open disk Ω in C under some natural conditions.

We note that for any $f \varepsilon$ H (D_r),

$$f(z) = \Sigma_{n=o}^{\infty} \frac{f^n(o)}{n!} z^n ,$$

by the Taylor's theorem. Thus the sequence $\{1, z, z^2, ...\}$ forms a basis for H (D_r) (See [1]). We want to generalize this feature of a basis in general topological algebras.

3.1 A subset $\{z^n\}_{n \geq o}$, $(z^o = e$, identity) of a topological algebra A is said to be a cyclic basis [5] or [10] if for each $x \varepsilon$ A there is a unique sequence $\{\alpha_i(x)\}_{i \geq 1}$ of scalars such that

$$x = \Sigma_{i=1}^{\infty} \alpha_i(x) z^i .$$

Thus, $1, z, z^2, ...$ is a cyclic basis of H (D_r), where z is the identity function of D_r .

It is worthwhile to note that the product of x, y in a topological algebra with a cyclic basis is given by the Cauchy product. Specifically, if $x = \Sigma_{i=1}^{\infty} \alpha_i z^i$ and $y = \Sigma_{i=o}^{\infty} \beta_i z^i$, then

$$xy = \Sigma_{n=o}^{\infty} (\Sigma_{k=o}^{n} \alpha_{n-k} \beta_k) z^n .$$

We elucidate some properties of a Frechet algebra A with a cyclic basis $\{z^n\}_{n \geq o}$.

3.2 (*a*) z cannot be invertible.

For, if so, then $z^{-1} = \Sigma_{n=o}^{\infty} \alpha_n z^n$ and so

$$e = zz^{-1} = \Sigma_{n=o}^{\infty} \alpha_n z^{n+1}$$

which contradicts the uniqueness of basis representation of e.

(*b*) For any complex λ, $z-\lambda e$ is invertable should be iff the series $\Sigma_{n=o}^{\infty}\left(\frac{z}{\lambda}\right)^n$ is convergent in A. And the inverse of $z-\lambda e$ is given by the formula :

$$(z-\lambda e)^{-1} = -\lambda^{-1}\,\Sigma_{n=o}^{\infty}\left(\frac{z}{\lambda}\right)^n .$$

To prove (*b*), suppose $(z-\lambda e)^{-1} = \Sigma_{n=o}^{\infty}\alpha_n z^n$.

Then $e = (z-\lambda e)\,\Sigma_{n=o}^{\infty}\alpha_n z^n = -\lambda\alpha_o e + \Sigma_{n=o}^{\infty}(\alpha_n - \lambda\alpha_{n+1})\,z^{n+1}$

and so we have : $-\lambda\alpha_o = 1$ and $\alpha_n - \lambda\alpha_{n+1} = 0$ for all $n \geq o$.

Since z is not invertible, $\lambda \neq 0$ and so $\alpha_o = -\lambda^{-1}$ and

$\alpha_{n+1} = \lambda^{-1}\alpha_n$ for $n \geq C$. By successive substitutions, the latter equation yields : $\alpha_n = -\lambda^{-(n+1)}$ for $n \geq o$. Thus the series $\Sigma_{n=a}^{\infty}\left(\frac{z}{\lambda}\right)^n$ is convergent. The rest follows easily.

(*c*) If $\sigma(z)$ denotes the spectrum of z, then for any complex number $\lambda \notin \sigma(z)$, $|\lambda| \geq \rho(z)$, the spectral radius of z.

For its proof, we see that $\sigma(z) = \{f(z) : f \varepsilon M(A)\}$ where M (A) is the maximal ideal space of A, which is identified with the set of all non-zero continuous multiplicative linear functionals on A. For each $f \varepsilon M(A)$,

$$f(z-\lambda e)^{-1} = -\lambda^{-1}\,\Sigma_{n=o}^{\infty}\left(\frac{f(z)}{\lambda}\right)^n$$

is convergent and so $\left|\frac{f(z)}{\lambda}\right| < 1$ for all $f \varepsilon M(A)$.

Thus

$$\rho(z) = \text{Sup}\,\{|f(z)| : f \varepsilon M(A)\} \leq |\lambda| .$$

3.3 Clearly, 3.2 (*c*) implies that if $r = \rho(z)$, $D_r \subset \sigma(z) \subset \overline{D_r}$ (the closure of the open disk D_r of radius r and centre o).

3.4 Now from 3.2 we can define a map $\phi : M(A) \to \sigma(z)$ by $\phi(f) = f(z)$ for all $f \varepsilon M(A)$. Clearly, ϕ is bijective and continuous when M (A) is given the Gelfand topology.

3.5 (*a*) If V is a subset of M (A) such that ϕ (V) has a limit point, then for any $x \varepsilon A$, $f(x) = 0$ for all $f \varepsilon V$ implies $x = 0$. Moreover, A is semi simple iff sgould be iff $\rho(x) > 0$.

For, if $x = \Sigma_{i=o}^{\infty} \alpha_i(x) z^i$, then for any $f \varepsilon M(A)$,

$f(x) = \Sigma_{i=o}^{\infty} \alpha_i(x) [f(z)]^i$ which is a power series converging to zero on the set ϕ (V) having a limit point and so by the identity theorem of Complex Analysis $\alpha_i(x) = 0$ for all $i \geq 1$. Therefore $x = 0$.

Further, if $\rho(z) = 0$ then $f(z) = 0$ for all $f \varepsilon M(A)$. Moreover, $f = \alpha_o$ [5] and so the radical, Rad A $\neq \{0\}$ *i.e.* A is not semisimple. On the other hand, if $x \varepsilon$ Rad A, then $f(x) = 0$ for all $f \varepsilon M(A)$. Since $D_{\rho(z)} \subset \sigma(z)$, $\rho(z) > 0$ implies by the first part that $x = 0$. Hence A is semisimple.

(*b*) $\sigma(z)$ is homeomorphic with M (A), provided $\sigma(z)$ is open.

Put $\tilde{x}(t) = \hat{x}(\phi^{-1}(t))$, where $\hat{x}$ is the Gelfand transform of x and ϕ defined in 3.4. Since ϕ is bijective and continuous, we have

$\tilde{x}(t) = \Sigma_{i=o}^{\infty} \alpha_i(x) \hat{z}^i(\phi^{-1}(t)) = \Sigma_{i=1}^{\infty} \alpha_i(x) t^n$ which represents an analytic function of $t \varepsilon \sigma(z)$. Hence $\tilde{x}$ is continuous on $\sigma(z)$ and this gives the continuity of ϕ^{-1} because the Gelfand topology is the weakest topology making each $\hat{x}$ continuous. From this it follows that $\sigma(z)$ is homeomorphic with M (A).

Using 3.5 (*a*), (*b*) we prove :

3.6 **Theorem** [10] A Frechet algebra A is isomorphic and homeomorphic with $H(\Omega)$ for some open disk $\Omega \subset C$ iff A has a cyclic basis $\{z^n\}_{n \geq o}$ with $\sigma(z)$ open. If $\rho(z) = \infty$, then A is isomorphic and homeomorphic with the Frechet algebra E of all entire functions.

Proof

Indeed, the "only if" part being trivial, we assume that A has a cyclic basis $\{z^n\}_{n \geq o}$ with $\sigma(z)$ open. If $\hat{x}$ denotes the Gelfand transform of $x \varepsilon$ A, we define $\tilde{x} = \hat{x} \, o \, \phi^{-1}$ where ϕ is defined in 3.4. Using 3.5 (b), we see that $\tilde{A} = \{ \tilde{x} : x \varepsilon A \}$ is a subalgebra of H (σ (z)), consisting of functions $\tilde{x}$ and equipped with the compact-open topology. If we define $\psi : \hat{A} \rightarrow \tilde{A}$ by $\psi(\hat{x}) = \tilde{x}$, in which $\hat{A} = \{\hat{x} : x \varepsilon A\}$, then using 3.5 ($b$) again, we see that ψ is an algebraic and topological isomorphism. If $g : A \rightarrow \hat{A}$ is the Gelfand map, then we want to show that $\psi \, og$ is the desired isomorphism on H (σ (z)). Since $\sigma(z)$ is open (hence $\rho(z) > 0$). A is semisimple (by 3.5 (a)) and so g and hence $\psi \, og$ is injective. Now if $f \varepsilon$ H(σ (z)) then by the usual functional calculus for Frechet algebras [9], there is $y \varepsilon$ A with $\hat{y}(\phi) = f(\hat{z}(\phi))$ for $\phi \varepsilon$ M (A). Equivalently, $f(t) = \hat{y}(\phi(t))$ for $t \varepsilon \sigma(z)$. Thus $f = \hat{y}$ and this proves that $\psi \, og$ is onto H (σ (z)). Furthermore, g as well as ψ being continuous implies $\psi \, og$ is continuous. Now the continuous linear map $\psi \, og$ of a Frechet algebra A onto the Frechet space H (σ (z)) is open by the Open Mapping theorem. The other part being similar, this completes the proof.

Note that the assumption that the spectrum $\sigma(z)$ be open in the above theorem can be weakened if we consider unconditional cyclic basis. For instance, we prove the following (see [10]). For completeness, we give a proof here.

3.7 Theorem

If $\{z^n\}_{n \geq o}$, is an unconditional cyclic basis of a Frechet algebra A, then $\sigma(z)$ is either open or compact. Moreover, M (A) is homeomorphic with $\sigma(z)$.

Proof

As shown above, $D_\rho(z) \subset \sigma(z) \subset \overline{D}_\rho(z)$. If $\sigma(z)$ contains a boundry point u, say, then for each $x \in A$, the series $\Sigma_{n=o}^{\infty} \alpha_i(x) u^i$ converges absolutely because the basis is unconditional. Hence if $w \in \mathbf{C}$ is such that $|w| = |u|$, then the convergence of the series $\Sigma_{i=1}^{\infty} \alpha_i(x) w^i$ implies that $f(z) = w$ is a multiplicative linear functional. Hence $w \in \sigma(z)$ ([9], Lemma 6.1 (a)) and so the disk $\sigma(z)$ contains all its boundry point. This proves the first part. For the second part, if $\sigma(z)$ is open then it is 3.5 (b).

This leads to the following, the proof of which is similar to that of 3.6.

3 8 Theorem

Let A be a Banach algebra in which the spectral norm ρ is equivalent to its given norm. (In other words, A is a Function algebra). If A has an unconditional cyclic basis $\{z^n\}_{n \geq o}$, then A is isomorphic to the algebra A (D_r) ($r = \rho(z)$) of all analytic functions on the disk D_r with the norm :

$$\|f\|_r = \Sigma_{n=1}^{\infty} \left| \frac{f^n(0)}{n!} \right| r^n < \infty .$$

We end with a problem.

3.9 Open Problem.

Let A be a Fre'chet algebra with identity and a basis $\{z_1^{n_1}, \ldots, z_m^{n_m}\}$, where $n_1, \ldots, n_m$ run over natural numbers, *i.e.* the algebra of polynomials in m variables is dense in A. When is A isomorphic and homeomorphic with H (Ω^m), the algebra of all holomorphic functions in m variables on a suitable open domain Ω^m in $\mathbf{C}^m$?

REFERENCES

1. Arsov, M.G. ; Proper bases and linear homeomorphism in spaces of analytic functions. *Math. Annalen,* 135 (1958), 235-243.
2. Day, M.M. ; Normed linear spaces. *Springer Verlag,* New ed. (1962).
3. Enflo, P. ; A counter example to the approximation problem in Banach spaces. *Acta Math.*, 130 (1973), 309-317.
4. Husain, T. and Liang, J. ; Multiplicative functionals on Frechet algebras with bases. *Canadian Jr. of Math.*, 29 (1977), 270-276.
5. Husain, T. and Watson, S.; Topological algebras with orthogonal Schauder bases. *Pacific Journal of Math.*, 91 (1980), 339-347.
6. Husain, T. and Watson, S. ; Algebras with unconditional orthogonal bases. *Proceedings, AMS,* 79 (1980), 539-545.
7. Husain, T. ; Multiplicative linear functionals on topological algebras. ***Pitman Research Notes, No. 85 Advanced Publishing Program, Pitman, London,*** Boston 1983.
8. Marti, J.T. ; Introduction to the Theory of bases. *Springer Verlag, Berlin,* 1969.
9. Michael, E.A. ; Locally multiplicatively-convex topological algebras. *Memoirs, AMS,* No. 11 (1952).
10. Savon, Z. and Watson, S. ; Metrizable algebras with cyclic bases. Preprint.
11. Singer, I. ; Bases in Banach spaces I. *Springer Verlag, Berlin,* 1970.
12. Zelazko, W. ; Banach algebras. *Elsevier, Amsterdam,* 1973.

The Punjab University
Journal of Mathematics
Vol. XIV—XV (1981-82), pp. 15-28

THE VECTOR SPACE OF MULTI-ADDITIVE ARITHMETICAL FUNCTIONS

By
L.M. CHAWLA

Department of Mathematics
Kansas State University,
Manhattan, Kansas 66506, U.S.A.

Abstract :

An arithmetical function $f(x_1, \dots, x_r)$, $r \geq 2$, with values in the field C of complex numbers is called here multi-additive, if $(x_i, y_i) = 1$, $i = 1, 2, \dots, r$ imply $f(x_1, \dots, x_i y_i, \dots, x_r)$ $= f(x_1, \dots, x_i, \dots, x_r) + f(x_1, \dots, y_i, \dots, x_r)$.

In this paper, we first prove that the set $\Omega(x_1, \dots, x_r)$ of all multi-additive functions is a vector space over C. Next for disjoint sets $\{x_{i_1}, \dots, x_{i_r}\}$, $\{x_{j_1}, \dots, x_{j_s}\}$ of distinct arguments we prove,

$$\Omega(x_{j_1}, \dots, x_{j_s}) \subseteq \text{Hom}\,(\Omega(x_{i_1}, \dots, x_{i_r}), \Omega(x_{i_1}, \dots, x_{i_r}, x_{j_1}, \dots, x_{j_s}))\,.$$

For a proper subset $\{i, j, \dots, k\}$ of the set $\{1, 2, \dots, r\}$, we construct certain H-homomorphisms $H_{i, j, \dots, k}$ of $\Omega(x_1, \dots, x_r)$ into $\Omega(x_i \cdot x_j, \dots, x_k)$ and prove that each $H_{i, j, \dots, k}$ can be decomposed uniquely to within order as the sum of certain irreducible H-homomorphisms and then develop a theory of H-homomorphs *i.e.* H-subspaces of $\Omega(x_i, x_j, \dots, x_k)$.

§ 1.

Definition 1. A real or complex valued arithmetical function $f(x_1, \ldots, x_r)$ of $r \geq 2$ arguments defined for all ordered lists $(x_1, \ldots, x_r)$ of positive integers x_i, $i = 1, 2, \ldots, r$, is said to be multi-additive if $(x_i, y_i) = 1$, $i = 1, 2, \ldots, r$ imply $f(x_1, \ldots, x_i y_i, \ldots, x_r)$
$= f(x_1, \ldots, x_i, \ldots, x_r) + f(x_1, \ldots, y_i, \ldots, x_r)$.

Since $f(x_1, \ldots, x_i \,.\, 1, \ldots, x_r) = f(x_1, \ldots, x_i, \ldots, x_r) + f(x_1, \ldots, 1, \ldots, x_r)$, it follows that $f(x_1, \ldots, 1, \ldots, x_r) = 0$, for all x_j, $j = 1, 2, \ldots, i-1, i+1, \ldots, r$, and hence $f(x_1, \ldots, x_r) = 0$ when 1 is substituted for one or more of thc arguments.

It is evident that the multi-additive arithmetical functions $f(x_1, \ldots, x_r)$ introduced here are different from the additive arithmetical functions of $r \geq 2$ arguments studied in [1] or in [3].

Let $\Omega(x_1, \ldots, x_r)$ denote the set of all multi-additive functions. In $\Omega(x_1, \ldots, x_r)$ define addition and multiplication by elements of the field C as usual. Thus, for any $f, g \varepsilon \Omega$ and $a \varepsilon C$, define

$$(f + g)(x_1, \ldots, x_r) = f(x_1, \ldots, x_r) + g(x_1, \ldots, x_r)$$

$$(a \,.\, f)(x_1, \ldots, x_r) = a f(x_1, \ldots, x_r).$$

Let $(x_i, y_i) = 1$, $i = 1, 2, \ldots, r$

and let $F(x_1, \ldots, x_r) = (f + g)(x_1, \ldots, x_r)$

$$= f(x_1, \ldots, x_r) + g(x_1, \ldots, x_r).$$

Then

$$F(x_1, \ldots, x_i y_i, \ldots, x_r) = f(x_1, \ldots, x_i y_i, \ldots, x_r) + g(x_1, \ldots, x_i y_i, \ldots, x_r).$$

By expanding and rearranging the terms on the R.H.S. of the above equation, we get

$F(x_1, \dots, x_i y_i, \dots, x_r) = F(x_1, \dots, x_i, \dots, x_r)$

$+ F(x_1, \dots, y_i, \dots, x_r)$. Thus, F is multi-additive and hence $F \varepsilon \Omega$.

Further the addition in Ω is easily proved to be associative and commutative. Finally define $(-f)(x_1, \dots, x_r) = -f(x_1, \dots, x_r)$ as the additive inverse of $f(x_1, \dots, x_r)$. Thus $\Omega(x_1, \dots, x_r)$ together with the addition as defined above is an additive abelian group with the zero element $0(x_1, \dots, x_r) = 0$ for all $(x_1, \dots, x_r)$.

By the definition of $(a \cdot f)(x_1, \dots x_r) = af(x_1, \dots, x_r)$, it follows that $a \cdot f$ is multi-additive and hence $a \cdot f \varepsilon \Omega$ and further $1 \cdot f = f$ and $a \cdot (f + g) = a \cdot f + a \cdot g$,

$(a + b) \cdot f = a \cdot f + b \cdot f$, for all $f, g \varepsilon \Omega$ and all $a, b \varepsilon C$. This completes the proof of

Theorem 1.1. The set $\Omega(x_1, \dots, x_r)$ of all multi-additive arithmetical functions is a vector space over C.

Examples of multi-additive arithmetical functions.

(1) Let $f_1(x_1), f_2(x_2), \dots, f_r(x_r)$, be any $r \geq 2$ given additive arithmetical functions respectively of the distinct arguments x_i, $i = 1, 2, \dots, r$. Consider the function formed by their product namely $f(x_1, \dots, x_r) = f_1(x_1) \cdot f_2(x_2), \dots, f_r(x_r)$. It is evident that $f(x_1, \dots, x_r)$ is multi-additive and hence $\varepsilon \Omega$. However, every multi-additive function in Ω need not be of the above form, since if $\quad f(x_1, \dots, x_r) = f_1(x_1) \cdot f_2(x_2) \cdot \dots \cdot f_r(x_r)$

and $\quad g(x_1, \dots, x_r) = g_1(x_1) \cdot g_2(x_2) \cdot \dots \cdot g_r(x_r)$

are in Ω, it is not true that $af + bg$ is necessarily a product of r distinct additive functions. However, the same conclusion is reached by the next example.

(2) Let $\theta(x_1, \dots, x_r)$ be any multi-multiplicative function [2], in the sense that $\theta \neq 0$, for any positive integral values of the arguments and further $(x_i, y_i) = 1$, $i = 1, 2, \dots, r$ imply

$$\theta(x_1, \dots, x_i y_i, \dots, x_r) = \theta(x_1, \dots, x_i, \dots, x_r) . \theta(x_1, \dots, y_i, \dots, x_r).$$

Then the function defined by $f(x_1, \dots, x_r) = \log \theta(x_1, \dots, x_r)$ is easily verified to be multi-additive.

Following the notation for the vector space $\Omega(x_1, \dots, x_r)$, let $\Omega(x_{i_1}, \dots, x_{i_r})$ and $\Omega(x_{j_1}, \dots, x_{j_s})$ denote respectively the vector spaces of all multi-additive functions $f(x_{i_1}, \dots, x_{i_r})$ and $g(x_{j_1}, \dots, x_{j_s})$, where the arguments involved are all distinct and the two sets $\{x_{i_1}, \dots, x_{i_r}\}$, $\{x_{j_1}, \dots, x_{j_s}\}$ are disjoint. Consider the function $p(x_{i_1}, \dots, x_{i_r}, \dots, x_{j_s})$ of the $r + s$ arguments defined by :

$$p(x_{i}, \dots, x_{i_r}, x_{j_i}, \dots, x_{j_s}) = f(x_{i_1}, \dots, x_{i_r}) . g(x_{j_1}, \dots, x_{j_s}) .$$

It is immediate that the function p is multi-additive and hence it belongs to the vector space $\Omega(x_{i_1}, \dots, x_{i_r}, x_{j_1}, \dots, x_{j_s})$. Let then $\Omega(x_{i_1}, \dots, x_{i_r}) . \Omega(x_{j_1}, \dots, x_{j_s})$ denote the set of all product functions p, for all $f \varepsilon \Omega(x_{i_1}, \dots, x_{i_r})$ and all $g \varepsilon \Omega(x_{j_1}, \dots, x_{j_s})$. We shall denote by

$\left[\Omega(x_{i_1}, \ldots, x_{i_r}) \cdot \Omega(x_{j_1}, \ldots, x_{j_s})\right]$, the subspace of $\Omega(x_{i_1}, \ldots, x_{i_r}, x_{j_1}, \ldots, x_{j_s})$ generated by the subset $\Omega(x_{i_1}, \ldots, x_{i_r}) \cdot \Omega(x_{j_1}, \ldots, x_{j_s})$. Note neither of the vector spaces $\Omega(x_{i_1}, \ldots, x_{i_r})$, $\Omega(x_{j_1}, \ldots, x_{j_s})$ is a subspace of $\Omega(x_{i_1}, \ldots, x_{i_r}, x_{j_1}, \ldots, x_{j_s})$ and nor in fact that of $\left[\Omega(x_{i_1}, \ldots, x_{i_r}) \cdot \Omega(x_{j_1}, \ldots, x_{j_s})\right]$.

We summarize the above conclusion in

Theorem 1.2 :

The product $\Omega(x_{i_1}, \ldots, x_{i_r}) \cdot \Omega(x_{j_1}, \ldots, x_{j_s})$ of the vector spaces $\Omega(x_{i_1}, \ldots, x_{i_r})$ and $\Omega(x_{j_1}, \ldots, x_{j_s})$ is a subset of the vector space $\Omega(x_{i_1}, \ldots, x_{i_r}, x_{j_1}, \ldots, x_{j_s})$ but neither of the vector spaces $\Omega(x_{i_1}, \ldots, x_{i_r})$, $\Omega(x_{j_1}, \ldots, x_{j_s})$ is a subspace of $\Omega(x_{i_1}, \ldots, x_{i_r}, x_{j_1}, \ldots, x_{j_s})$ nor that of $\left[\Omega(x_{i_1}, \ldots, x_{i_r}) \cdot \Omega(x_{j_1}, \ldots, x_{j_s})\right]$.

In contrast to Theorem 1.2, the vector spaces $\Omega(x_{i_1}, \ldots, x_{i_r})$, $\Omega(x_{j_1}, \ldots, x_{j_s})$ and $\Omega(x_{i_1}, \ldots, x_{i_r}, x_{j_1}, \ldots, x_{j_s})$ are related to each other in an important way as proved in

Theorem 1.3 :

(*a*) $\Omega(x_{j_1}, \ldots, x_{j_s}) \subseteq$

$$\text{Hom}\,(\Omega(x_{i_1}, \ldots, x_{i_r}), \Omega(x_{i_1}, \ldots, x_{i_r}, x_{j_1}, \ldots, x_{j_s}))$$

(*b*) $\Omega(x_{i_1}, \dots, x_{i_r}) \subseteq$

$$\text{Hom}\ (\Omega(x_{j_1}, \dots, x_{j_s}), \Omega(x_{i_1}, \dots, x_{i_r}, x_{j_1}, \dots, x_{j_s}))$$

To prove (*a*), let $f(x_{i_1}, \dots, x_{i_r})$ and $g(x_{i_1}, \dots, x_{i_r})$ $\varepsilon\, \Omega(x_{i_1}, \dots, x_{i_r})$ and let $h(x_{j_1}, \dots, x_{j_s})\ \varepsilon\ \Omega(x_{j_1}, \dots, x_{j_s})$, then

$$f(x_{i_1}, \dots, x_{i_r}) \,.\, h(x_{j_1}, \dots, x_{j_s}) = \overline{f}(x_{i_1}, \dots, x_{i_r}, x_{j_1}, \dots, x_{j_s})$$

is multi-additive and $\varepsilon\ \Omega(x_{i_1}, \dots, x_{i_r}, x_{j_1}, \dots, x_{j_s})$ (Theorem 1.2 above). Further for all $f, g\ \varepsilon\ \Omega(x_{i_1}, \dots, x_{i_r})$ and $a, b\ \varepsilon\ C$, we have

$$(a \,.\, f + b \,.\, g) \,.\, h = af \,.\, h + bg \,.\, h$$

$$= a\,\overline{f} + b\,\overline{g}\ \varepsilon\ \Omega(x_{i_1}, \dots, x_{i_r}, x_{j_1}, \dots, x_{j_s})\,.$$

It follows that each element h of $\Omega(x_{j_1}, \dots, x_{j_s})$ is a homomorphism of $\Omega(x_{i_1}, \dots, x_{i_r})$ into $\Omega(x_{i_1}, \dots, x_{i_r}, x_{j_1}, \dots, x_{j_s})$.

This proves (*a*) and then (*b*) follows by symmetry.

Corollary. For any vector space $\Omega(x_1, \dots, x_r)$ we have

$$\Omega(x_{j_1}, \dots, x_{j_l}) \subseteq \text{Hom}\ (\Omega(x_{i_1}, \dots, x_{i_k}), \Omega(x_1, \dots, x_r))$$

where $i_1, i_2, \dots, i_k;\ j_1, \dots, j_l$ is any partition of the r integers $1, 2, 3, \dots, r$ into two disjoint subsets.

Theorem 1.3 gives a certain class of homormorphisms of the vector space $\Omega(x_1, \dots, x_r)$ into the "higher" vector space $\Omega(x_1, \dots, x_r, x_{r+1}, \dots, x_{r+t})$. In the next section, we construct a class of homomorphisms of the vector space $\Omega(x_1, \dots, x_r)$ into the "lower" vector space $\Omega(x_i, x_j, \ldots, x_k)$ where $\{i, j, \ldots, k\}$ is a proper subset of $\{1, 2, 3, \ldots, r\}$.

§ **2.**

Let $\{i, j, \ldots, k\}$ be a proper subset of s integers of the set $\{1, 2, \ldots, r\}$, $1 \leq s < r$. For a given ordered tuple

$$\alpha = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{k-1}, a_{k+1}, \ldots, a_r)$$

of $r - s$ positive integers a's, define the mapping $H^{(\alpha)}_{i, j, \ldots, k}$ of the vector space $\Omega(x_1 \ldots, x_r)$ into the vector space $\Omega(x_i, x_j, \ldots, x_k)$ by

$$(i) \quad f(x_1, \ldots, x_r)\, H^{(\alpha)}_{i, j, \ldots, k}$$

$$= f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_{k-1}, x_k, a_{k+1}, \ldots, a_r)$$

$$= \overline{f}(x_i, x_j \ldots, x_k), \text{ say.}$$

$$(ii) \quad (a \,.\, f(x_1, \ldots, x_r))\, H^{(\alpha)}_{i, j, \ldots, k}$$

$$= a \,.\, (f(x_1, \ldots, x_r))\, H^{(\alpha)}_{i, j, \ldots, k}$$

for all $f \varepsilon \Omega(x_1, \ldots, x_r)$ and all $a \varepsilon C$.

It is easily verified that $\overline{f}(x_i, x_j, \ldots, x_k)$ is multi-additive and hence it $\varepsilon\, \Omega(x_i, x_j, \ldots, x_k)$ and further for any $f, g \,\varepsilon\, \Omega(x_1, \ldots, x_r)$ and $a, b \,\varepsilon\, C$, we have $(a \,.\, f + b \,.\, g)\, H^{(\alpha)}_{i, j, \ldots, k}$

$$= a \,.\, \overline{f} + b \,.\, \overline{g} \,\varepsilon\, \Omega(x_i, x_j, \ldots, x_k).$$

Thus, we have proved

Lemma 2.1 : The mapping $H^{(\alpha)}_{i, j, \ldots, k}$ is a homomorphism of the vector space $\Omega(x_1, \ldots, x_r)$ into the vector space $\Omega(x_i, x_j, \ldots, x_k)$.

Definition 2. The homomorphism $H^{(\alpha)}_{i, j, \ldots k}$ is called irreducible, if each coordinate of the tuple
$\alpha = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{k-1}, a_{k+1}, \ldots, a_r)$ is a prime or a power of a prime.

We denote the homomorph of $\Omega(x_1, \ldots, x_r)$ under the homomorphism $H^{(\alpha)}_{i, j, \ldots, k}$ by $\Omega(x_1, \ldots, x_r)\, H^{(\alpha)}_{i, j, \ldots, k}$ or explicitly by $\Omega(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots\ a_{k-1}, x_k, a_{k+1}, \ldots, a_r)$ as the situation may require. It is evident that each of the homomorphs $(\Omega)\ H^{(\alpha)}_{i, j, \ldots, k}$, for variable α's, is a subspace of the vector space $\Omega(x_i, x_j, \ldots, x_k)$. We call each such homomorph an H-subspace of $\Omega(x_1, x_j, \ldots, x_k)$. An H-subspace of $\Omega(x_i, x_j, \ldots, x_k)$ is then called irreducible H-subspace if it is the homomorph of $\Omega(x_1, \ldots, x_r)$ under an irreducible H-homomorphism $H^{(\alpha)}_{i, j, \ldots, k}$.

Theorem 2.2. Any H-homomorphism $H^{(\alpha)}_{i, j, \ldots, k}$ can be decomposed uniquely to within order as the sum of irreducible H-homomorphisms.

Let $\alpha = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{k-1}, a_{k+1}, \ldots, a_r)$,

then

$$(f(x_1 \ldots, x_r))\, H^{(\alpha)}_{i, j, \ldots, k}$$
$$= f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_{k-1}, x_k, a_{k+1}, \ldots, a_r)$$

$$= f(\Pi_{p_1}^{e_1}, \dots, \Pi_{p_{i-1}}^{e_{i-1}}, x_i, \Pi_{p_{i+1}}^{e_{i+1}}, \dots, \Pi_{p_{k-1}}^{e_{k-1}}, x_k, \Pi_{p_{k+1}}^{e_{k+1}}, \dots, \Pi_{p_r}^{e_r})$$

$$= \Sigma f(p_1^{e_1}, \dots, p_{i-1}^{e_{i-1}}, x_i, p_{i+1}^{e_{i+1}}, \dots, p_{k-1}^{e_{k-1}}, x_k, p_{k+1}^{e_{k+1}}, \dots, p_r^{e_r}),$$

the summation extending over all the greatest powers of primes which divide each of the a's in f. Thus

$$(f(x_1, \dots, x_r))\, H_{i,j,\dots,k}^{(\alpha)} = \Sigma f(x_1, \dots, x_r)\, H_{i,j,\dots,k}(p_1^{e_1}, \dots, p_{i-1}^{e_{i-1}}, p_{i+1}^{e_{i+1}}, \dots, p_r^{e_r}) \text{ for all } f(x_1, \dots, x_r).$$

Hence $H_{i,j \dots, k}^{(\alpha)} = \Sigma H_{i,j,\dots,k}^{(\gamma)}$, where each $H_{i,j,\dots,k}^{(\gamma)}$ is irreducible and this sum is uniquely determined to within order, because of unique representation of the a's in α as products of primes. This proves the theorem.

Let $\omega(n)$, as usual, denote the number of distinct primes which divide the positive integer n with $\omega(1) = 0$, we then have

Corollary. The number of irreducible H-homomorphisms in the sum of $H_{i,j,\dots,k}^{(\alpha)}$, where $\alpha = (a_1, a_2, \dots, a_{i-1}, a_{i+1}, \dots, a_{k-1}, a_{k+1}, \dots, a_r)$ is given by $W(\alpha) = \omega(a_1) \cdot \dots \cdot \omega(a_{i-1})\, \omega(a_{i+1}) \cdot \dots \cdot \omega(a_{k-1})\, \omega(a_{k+1}) \cdot \dots \cdot \omega(a_r)$.

Theorem 2.3. Each H-subspace

$\Omega(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_{k-1}, x_k, a_{k+1} \dots, a_r)$

of $\Omega(x_i, x_j, \dots, x_k)$ can be decomposed uniquely to within order as the sum of W (α) irreducible H-subspaces of $\Omega(x_i, x_j, \dots, x_k)$.

This follows directly from Theorem 2.2 and the corollary above. We need hardly point out that the converses of Theorems 2.2 and 2.3 are not true, since the sum of any two or more irreducible H-homomorphisms need not be an H-homomorphism.

Let $S_{i,j,\dots,k}\, \Omega(x_1, \dots, x_r)$ or briefly $S_{i,j,\dots,k}^{(\Omega)}$ denote the linear sum of all the H-subspaces $(\Omega(x_1, \dots, x_r))\, H_{i,j,\dots,k}^{(\alpha)}$ of $\Omega(x_i, x_j, \dots, x_k)$. One may characterize $S_{i,j,\dots,k}^{(\Omega)}$ by any one of the following ways.

(*i*) $S_{i,j,\dots,k}^{(\Omega)}$ is the subspace of $\Omega(x_i, x_j, \dots, x_k)$ generated by the logical union of all the H-subspaces of $\Omega(x_i, x_j, \dots, x_l)$

(*ii*) $S_{i,j,\dots,k}^{(\Omega)}$ is the intersection of all subspaces of $\Omega(x_i, x_j, \dots, x_k)$ which contains each of the H-subspaces of $\Omega(x_i, x_j, \dots, x_k)$, or in fact, by Theorem 2.3, it is the intersection of all subspaces of $\Omega(x_i, x_j, \dots, x_k)$ which contains each of its irreducible H-subspaces. Thus we have

Lemma 2.4. Every element of $S_{i\,j,\dots,k}^{(\Omega)}$ is either of the form

$$f(p_1^{e_1}, \dots, p_{i-1}^{e_{i-1}}, x_i, p_{i+1}^{e_{i+1}}, \dots, p_{k-1}^{e_{k-1}}, x_k, p_{k+1}^{e_{k+1}}, \dots, p_r^{e_r})$$

or is a linear sum of such functions.

Let $S_{i_1, i_2, \dots, i_k}\,\Omega(x_1, \dots, x_r)$ and $S_{j_1, \dots, j_l}\,\Omega(x_1, \dots, x_r)$ be respectively the linear sums of all the H-subspaces of $\Omega(x_{i_1}, \dots, x_{i_k})$ and $\Omega(x_{j_1}, \dots, x_{j_l})$ where $\{i_1, \dots, i_k\}, \{j_1, \dots, j_l\}$ are proper disjoint subsets of $\{1, 2, \dots, r\}$. Then by Theorem 1.2, $\left[S^{(\Omega)}_{i_1, \dots, i_k} \cdot S^{(\Omega)}_{j_1, \dots, j_l} \right]$ is a subspace of $\Omega(x_{i_1}, \dots, x_{i_k}, x_{j_1}, \dots, x_{j_l})$ but neither of the subspaces $S^{(\Omega)}_{i_1, \dots, i_k}, S^{(\Omega)}_{j_1, \dots, j_l}$ is a subspace of $\Omega(x_{i_1}, \dots, x_{i_k}, x_{j_1}, \dots, x_{j_l})$ nor in fact that of $\left[S^{(\Omega)}_{i_1, \dots, i_k} \cdot S^{(\Omega)}_{j_1, \dots, j_l} \right]$. However, they are related to each other and to the subspace generated by their product in a special way as proved in the next Theorem, for which we need the following.

Lemma 2.5.

$$S_{e_1, \dots, e_l}(S^{(\Omega)}_{i_1, \dots, i_k}) = S^{(\Omega)}_{e_1, \dots, e_l}$$ where $\{e_1, e_2, \dots, e_l\}$ is a proper subset of the set $\{i_1, \dots, i_k\}$.

By Lemma 2.4, any element of $S^{(\Omega)}_{i_1, \dots, i_1}$ is either of the form

$$f(p_1^{e_1}, \dots, p_{i_1-1}^{e_{i_1}-1}, x_{i_1}, p_{i_1+1}^{e_{i_1}+1}, \dots, p_{i_k-1}^{e_{i_k}-1}, x_{i_k}, p_{i_k+1}^{e_{i_k}+1}, \dots, p_r^{e_r})$$

or is a linear combination of elements of this form. Hence

$$S_{e_1, \dots, e_l}(S^{(\Omega)}_{i_1, \dots, i_k}) \subseteq S^{(\Omega)}_{e_1, \dots, e_l}$$

The reverse inclusion follows by the definition of $S^{(\Omega)}_{e_1, \dots, e_l}$. This proves the lemma.

Theorem 2.6.

(*a*) $$S_{i_1, \dots, i_k}\left(\left[S^{(\Omega)}_{i_1, \dots, i_k} \cdot S^{(\Omega)}_{j_1, \dots, j_l} \right]\right)$$

$$= S_{i_1, \dots, i_k}(S^{(\Omega)}_{i_1, \dots, i_k, j_1, \dots, j_l}) = S^{(\Omega)}_{i_1, i_2, \dots, i_k}$$

(*b*) $$S_{j_1, \dots, j_l}\left(\left[S^{(\Omega)}_{i_1, \dots, i_k} \cdot S^{(\Omega)}_{j_1, \dots, j_l} \right]\right)$$

$$= S_{j_1, \dots, j_l}(S^{(\Omega)}_{i_1, \dots, i_k, j_1, \dots, j_l}) = S^{(\Omega)}_{j_1, j_2, \dots, j_l}$$

The second equality in both (*a*) and (*b*) follows directly from Lemma 2.5. To prove the first equality in (*a*), consider any element of the subspace $\left[S^{(\Omega)}_{i_1, \dots, i_k} \cdot S^{(\Omega)}_{j_1, \dots, j_l} \right]$, it is of the form $\Sigma \lambda_t f_t g_t$, where each $f_t \in S^{(\Omega)}_{i_1, \dots, i_k}$ and each $g_t \in S^{(\Omega)}_{j_1, \dots, j_l}$ and $\lambda_t \in C$. Hence by definition of $S_{i_1, \dots, i_k}$, any element of $S_{i_1, \dots, i_k}\left(\left[S^{(\Omega)}_{i_1, \dots, i_k} \cdot S^{(\Omega)}_{j_1, \dots, j_l} \right]\right)$ is of the form $\Sigma \mu_t f_t$, where $\mu_t \in C$ and $f_t \in S^{(\Omega)}_{i_1, \dots, i_k}$. Thus

$$S_{i_1, \dots, i_k}\left(\left[S^{(\Omega)}_{i_1, \dots, i_k} \cdot S^{(\Omega)}_{j_1, \dots, j_l} \right]\right) \subseteq S^{(\Omega)}_{i_1, \dots, i_k}.$$

Conversely taking any element of $S^{(\Omega)}_{i_1, \dots, i_k}$ in the form $\Sigma \lambda_t f_t$ where $\lambda_t \in C$ and

$f_t \in S_{i_1, \dots, i_k}^{(\Omega)}$ and then writing each $\lambda_t = c_t \cdot \mu_t$ where μ_t is the value of a given g_t in $S_{j_1, \dots, j_l}^{(\Omega)}$ by replacing each $x_{j_1}, \dots, x_{j_l}$ by powers of any given primes, it follows that $\Sigma \lambda_t f_t = \Sigma c_t \mu_t f_t$ is in $S_{i_1, \dots, i_k}\left(\left[S_{i_1, \dots, i_k}^{(\Omega)} \cdot S_{j_1, \dots, j_l}^{(\Omega)} \right] \right)$,

and hence the reverse inclusion holds. This proves (*a*) completely. Then (*b*) follows by symmetry.

Corollary.

(*a*) $S_i \left(\left[S_i(\Omega) \cdot S_j(\Omega) \right] \right) = S_i \left(S_{i,j}(\Omega) \right) = S_i(\Omega)$

(*b*) $S_{j,k} \left(\left[S_i(\Omega) \cdot S_{j,k}(\Omega) \right] \right) = S_{j,k} \left(S_{i,j,k}(\Omega) \right)$

$$= S_{j,k}(\Omega) .$$

In conclusion, I would like to add that the properties of the vector space $\Omega(x_1, \dots, x_r)$ established in Theorems 1.2, 1.3 and lemma 2.1 hold for the parent vector space of all arithmetic functions $f(x_1, \dots, x_r)$, $r \geq 2$ as well, but they are not true for every subspace of the parent vector space. For instance, these properties are not inherited by the subspace of all additive arithmetic functions, $f(x_1, \dots, x_r)$, $r \geq 2$.

Further the properties of $\Omega(x_1, \dots, x_r)$ proved in Theorems 2.2, 2.3, 2.4, 2.5 and 2.6 are neither shared by the parent vector space nor by the subspace of all additive arithmetic functions.

REFERENCES

1. Chawla, L.M., Distributive Arithmetical functions of two or more variables (to appear).
2. Chawla, L.M., The Groups of multi-multiplicative arithmetical functions (to appear).
3. Muclaire, Jean-Loup, Quelques results sur les fonctions additives a deun variables, C.R. *Acad. Sci. Paris*, *Ser* A-B, **274** (1972) **A** 373-A 376 and M.R. Vol. 45 No. 1859.

The Punjab University
Journal of Mathematics
Vol. XIV—XV 1981-82), pp. 29-40

SOME COMMENTS ON CATASTROPHE THEORY

By

B.A. SALEEMI*

Faculty of Science, King Abdulaziz University,
Jeddah, Saudi Arabia

I. Introduction.

According to Thom [6, p. 4], the catastrophe theory is the qualitative study of the natural processes occuring in the space-time manifold R^4 via their postulated 'differential models'. A *model*, as defined in [1], is an abstract. simplified, mathematical construct, related to a part of reality and created for a particular purpose. For instance, if P is a natural process and M is its postulated differential model, then a study of the properties of M can enable us to draw qualitative conclusions about the nature of singularities of P. In explaining the philosophy behind the catastrophe theory, it is claimed in [3, 6, 7] that the only part of a natural process which can be observed, is its 'catastrophic set'—a philosophically plausible speculation because of the complementary relationship between the observable and non-observable parts but mathematically it is an unproven assertion. It is not unreasonable to surmise that a catastrophic set may itself have its own catastrophic set and so on. An analogous situation is the process of peeling off an onion ; and it is also probably related in 'some way' to the concept of stratification

* The author wishes to express his gratitude to Professor E.C. Zeeman and the referee for their helpful comments and suggestions.

of manifolds [3]. To circumvent this rather recondite and ticklish problem, Thom [6, 1 . 3] asserts that our local models do not imply about the "ultimate nature of reality". I, for one, would like to make a frank confession that I do not know whether the aforesaid feature of models be recognised as a merit or a limitation *vis-a-vis* their applicability to the study of natural laws.

In [6, 7] it is unambiguously asserted that the catastrophe theory is a 'method' which is amply powerful to provide an explanation why continuous causes give rise to discontinuous effects. For example, the phenomenon of shock waves [4 (2)] which is of paramount importance in engineering and physics or the phenomenon of aggression [7] in psychology can be explained by using catastrophe theory. Presumably it would have been more interesting if one could have been able to use 'continuous models' instead of differential models to give rational justification for the occurrence of discontinuities during a continuous process. One common reason for choosing differential models given in [3, 6, 7] is that all natural processes are supposedly governed by their 'gradient dynamics' and hence the only mathematically justifiable practical approach to model nature is to postulate differential models. This argument seems sound because almost all known physical processes occuring in R^4 are governed by potential functions and, as a consequence, have their associated dynamics with them. However one insurmountable difficulty inherent in the theory of models is the amount of "inexactness" involved in the choice of governing potential funcuons. To give an appraisal of the effect of this 'uninhibited choice' of potential functions which govern a particular process, we cite some very familiar and elementary cases in the sequel.

2. Algebraic Equations.

Let C and X be two planes with coordinates c, b and x_1, x_2 respectively. Let $\phi : X \longrightarrow C$ be defined by :

$$c = x_1 x_2 , \qquad (1)$$

$$b = -\tfrac{1}{2}(x_1 + x_2) .$$

Then the Jacobian matrix of ϕ, namely, ϕ_*, is given by

$$\phi_* = \begin{bmatrix} x_2 & x_1 \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \qquad (2)$$

The matrix ϕ_* has rank 2 except at the points along the line $x_1 = x_2$ (See Figs (i), (ii)).

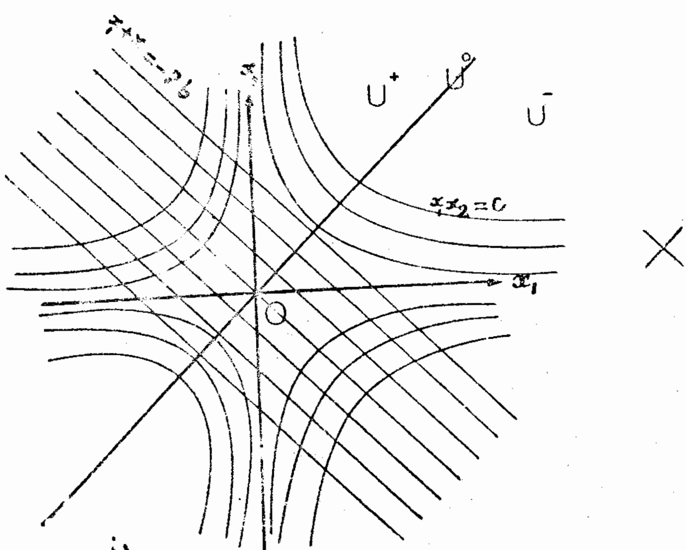


Fig. (i)

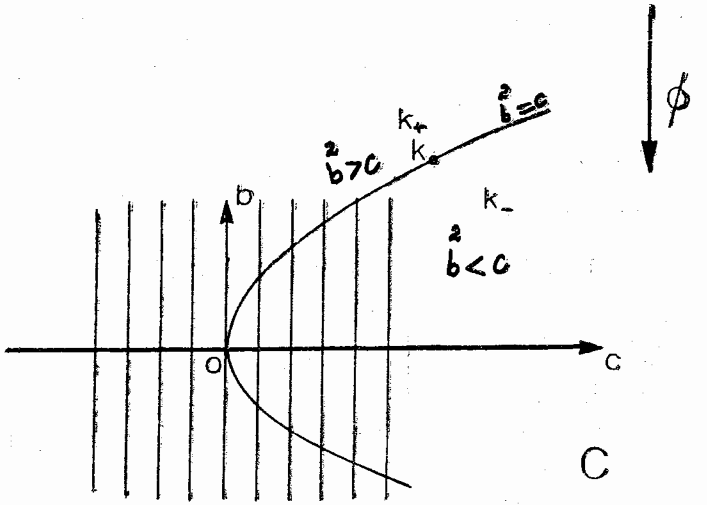


Fig. (*ii*)

The line $x_1 = x_2$ corresponds under ϕ to the parabola $b^2 = c$ in the C-plane. We may write

$$C = K_- \cup K_o \cup K_+ ,$$

where

$$K_- = \{ (c, b) \in C : b^2 < c \},$$

$$K_o = \{ (c, b) \in C : b^2 = c \},$$

and $K_+ = \{ (c, b) \in C : b^2 > c \}$,

The set $C\text{-}K_o$ has two components K_- and K_+ which are both open in C. Observing that the points (x_1 , x_2) and (x_2 , x_1) have the same image in C, we conclude that ϕ is *not* globally injective. Let $U_+ = \{ (x_1 , x_2) : x_2 > x_1 \}$, and $U_- = \{ (x_1 , x_2) : x_2 < x_1 \}$. Then ϕ/U_+ and ϕ / U_- are both injective maps.

Since

$$4(b^2 - c) = (x_1 - x_2)^2, \tag{3}$$

it follows that

$$\phi / U_+ : U_+ \longrightarrow K_+ \text{ and } \phi / U_- : U_- \longrightarrow K_+$$

are both diffeomorphisms. Hence no point of X under ϕ has its image in K_-. It may also be noted that $\phi / X - \phi^{-1}(K_o)$ is an immersion.

To give an alternative interpretation of the behaviour of $\phi : X \longrightarrow C$, we may think that every quadratic equation

$$x^2 + 2bx + c = 0 \tag{4}$$

is represented by a point $(c, b) \in C$. Let

$$\triangle = b^2 - c \tag{5}$$

be the discriminant of equation (4). Then the quadratic equations represented by the points of K_- *i.e* ; $\triangle < 0$, have no real roots. For equations represented in K_- we have

$$x_1 = -b + i(-\triangle)^{\frac{1}{2}}, \quad i = \sqrt{-1},$$

$$x_2 = -b - i(-\triangle)^{\frac{1}{2}}; \tag{6}$$

and hence

$$b = -\tfrac{1}{2}(x_1 + x_2),$$

$$c = x_1 x_2,$$

Now $b^2 < c$ implies that

$$(x_1 - x_2)^2 < 0, \tag{7}$$

which is impossible for real numbers x_1, x_2.

Remarks.

1. The function $\phi : X \longrightarrow C$ does not map any point into K_- without telling us why. Further the definition of ϕ remains unaffected whether x_1, x_2 are real or a pair of complex conjugate numbers.

2. The quadratic equation (4) defines x as a two-valued function on C. These two values x_1, x_2 are :

(*i*) real and distinct on K_+
(*ii*) Real and coincident on K_o
(*iii*) complex conjugates on K_- .

Let Γ_θ be the intersection of the loci

$$x_1 = -b + \sqrt{\triangle} \text{ and } x_2 = -b - \sqrt{\triangle} . \qquad (8)$$

Then Γ_θ does not depend upon the nature of $\triangle$ and always projects onto K_o . A routine computation shows that, in general, the Gaussian curvatures of the surfaces

$$x_1 = -b + \sqrt{\triangle}$$

and

$$x_2 = -b - \sqrt{\triangle}$$

are unequal. However when we try to approach U_o (*i.e.* when $(b^2 - c) \rightarrow 0$) along these surfaces, the Gaussian curvatures of both the surfaces tend to infinity. This observation, though apparently simple and trivial, is of great consequence because it brings to surface a hitherto unsuspected relationship between curvature and the catastrophe theory.

To assess the impact of the arbitrariness involved in the choice of potential functions on the final outcomes, we now apply the above analysis to the Example 1 [7, p. 3] pertaining to the dogs behaviour under varying circumstances. Let

$x =$ behaviour variable,
$c =$ fear,

and

$b =$ rage.

The variable x is usually called the 'state variable' or 'internal variable' and c, b are known as 'control parameters' or 'control variables'. Let

$$V(x, c, b) = \frac{x^3}{3} + bx^2 + cx$$

be the potential function governing the behaviour of the dog. Then the surface of equilibria M is given by

$$\frac{dV}{dx} = x^2 + 2bx + c = 0 ;$$

and the singularity set is the subset of M represented by

$$\frac{d^2V}{dx^2} = 2x + 2b = 0.$$

Hence the bifurcation set B, obtained by eliminating x from $\frac{dV}{dx} = 0$

and $\frac{d^2V}{dx^2} = 0$, is precisely the discriminant

$$\Delta = b^2 - c = 0.$$

The two sheets $x_1 = -b + \sqrt{b^2 - c}$

and

$$x_2 = -b - \sqrt{b^2 - c}$$

fold along Γ_0 and the catastrophe is the fold-catastrophe. Along U_0 (*i.e.* $x_1 = x_2$), the dog can change his mood from attack to flight or *vice versa* without being detected (see Fig. (*iii*) (*a*)).

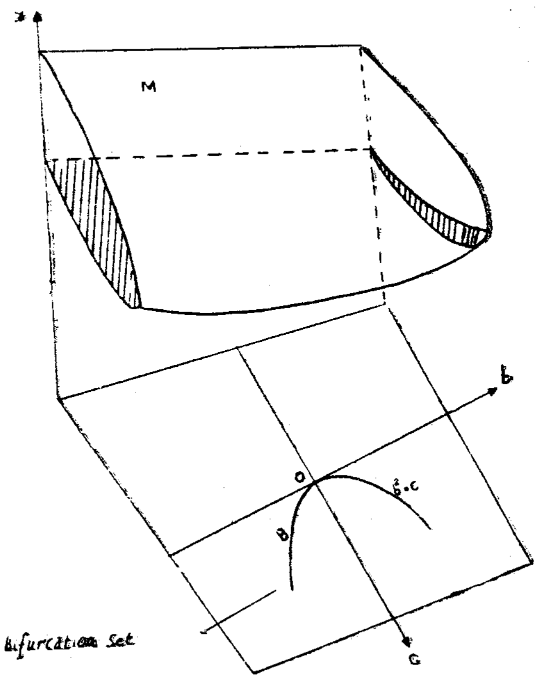


Fig (iii) (a)

Let us now postulate that the dog's behaviour is governed by the potential function

$$V(x, c, b) = \frac{x^4}{4} + \frac{bx^2}{2} + cx. \quad (8)$$

Then M is given by

$$x^3 + bx + c = 0 \quad (8')$$

and the singularity set (a subset of M) is

$$3x^2 + b = 0. \quad (8'')$$

Elimination of x from (8′) and (8″) yields that

$$\Delta = 4b^3 + 27c^2 = 0$$

is the bifurcation set and the catastrophe is the cusp-catastrophe (see Fig. (*iii*) (a') and Fig. (*iii*) (a'') as given in [7].

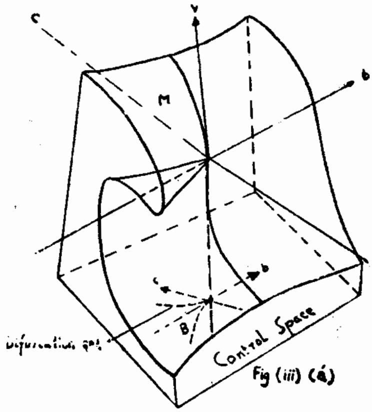


Fig (iii) (a')

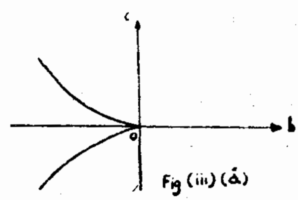
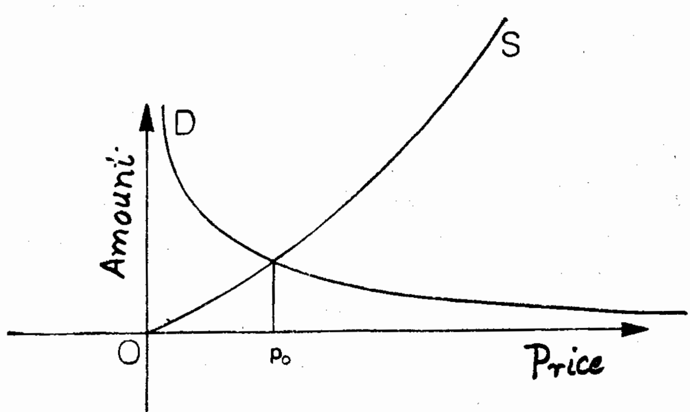


Fig (iii) (a'')

Thus we see that the 'postulated model' introduces drastic changes in our final conclusions.

3. **Supply and Demand Analysis.**

Let S and D be the supply and demand diagrams (see Fig. (*iv*)) [1, 4].



Fig(iv)

The price p_o is called the equilibrium price where the curves S and D intersect. Let us postulate that the supply s, demand d and price p are connected by :

$$s = bp + cp^2 \text{ and } d = \frac{k}{p} . \tag{10}$$

Then the equilibrium price p_o is given by

$$s = d$$

or

$$cp^3 + bp^2 - k = 0. \tag{11}$$

Starting from s_1 and following the arrows, we obtain the following two sequences $\{ p_{2n} \}$ and $\{ p_{2n+1} \}$ with (See Fig. (v))

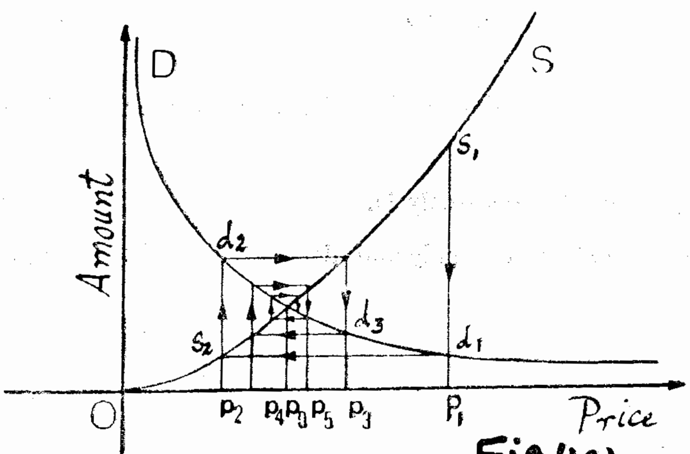


Fig (v)

$$\lim_{n\to\infty} p_{2n} = p_o = \lim_{n\to\infty} (p_{2n+1}) \tag{12}$$

Thus the supply and demand adjust by the passage of time towards p_o . Further the bifurcation set corresponding to

$$V(p, c, b) = \frac{c}{4} p^4 + \frac{b}{3} p^3 - kp$$

is

$$\triangle = ck^3 (4b^3 + 27c^2) = 0. \tag{13}$$

Hence the catastrophe is a cusp-catastrophe as $\triangle = 0$ gives

$$4b^3 = -27c^2$$

If we postulate that the supply-demand process is governed by the potential function $V(p, c, b)$ with

$$\triangle = b - c^4 + 2c^2 , \qquad (14)$$

then there will be two absolute minima (see Fig. (*vi*)) and consequently 'economical explosion' will take place [6]. Thus the unwelcome factor of 'personal bias'

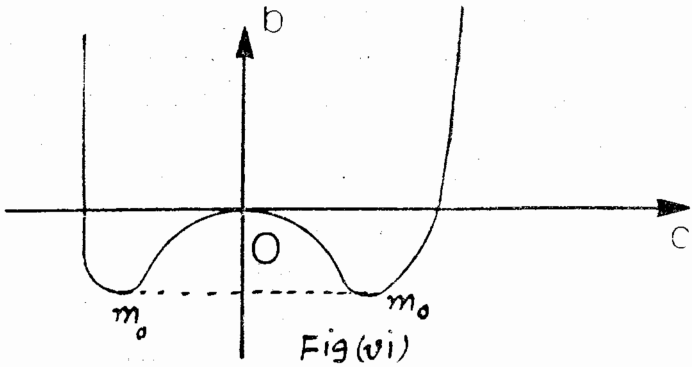


Fig (vi)

in the choice of potential functions affects our conclusions drastically. Probably something is still lacking in our comprehension of postulated differential models because the amount of subjectivity present in the choice of models makes our studies a little dubious.

REFERENCES

1. E.A. Bender, Introduction to Mathematical Modeling John Wiley & Sons, 1978.

2. G.B. Kolata, 'Prince Without Clothes Science, Vol. 196, No. 4287, Page 287, 15 April, 1977.

3. Y.C. Lu, Singularity Theory, Universitext Springer-Verlag, 1976.

4. J. Von Neumann and O. Morgenstern 1. The Theory of Games and Economic Behaviour, Princeton, 1944.
   2. Collected Works, Vol. VI (Page 176).

5. T. Poston and E.A.R. Woodcock, A Geometrical Study of Elementary Catastrophes,
   Lecture Notes in Math. No. 373, Springer-Verlag, 1974.

6. R. Thom, Structural Stability and Morphogenesis. (English Transl. by D.H. Fowler) Benjamin, New York, 1975.

7. E.C. Zeeman: Catastrophe Theory, Scientific American, Vol. 234, No. 4, 1976.

8. E.C. Zeeman and C.A. Isnard, Use of Models in the Social Sciences, Tavistack Publications, London, 1976.

The Punjab University
Journal of Mathematics
Vol. XIV—XV (1981-82), pp. 41-50

UNIFICATION OF ELECTROMAGNETISM WITH STRONG INTERACTIONS

By

FAYYAZUDDIN
Physics Department
King Saud University
Riyadh, Saudi Arabia

and

RIAZUDDIN
Physics Department
University of Petroleum & Minerals
Dhahran, Saudi Arabia

Abstract

Unification of Quantum Electrodynamics (QED) with Quantum Chromodynamics (QCD) based on local gauge group SU_c (4) is considered. Some consequences of this unification are examined.

It is believed that quarks carry colour. The interactions between coloured quarks are mediated by massless vector coloured gluons. The underlying theory for these interactions is a gauge theory called quantum chromodynamics (QCD). It is believed that QCD is the best candidate for a theory of the strong interactions. Each quark carries three colours. Thus the colour gauge group is SU_c (3) and QCD has a non-Abelian character in contrast to quantum electrodynamics (QED) which is Abelian.

On the other hand, one sees a remarkable similarity between QED and QCD. Both belong to local symmetry groups which are exact. It is, therefore, natural to consider the unification of electromagnetic interaction with strong interactions given by QCD. Both photon and eight coloured vector gluons being mediators of electromagnetic and strong interactions [which are given by exact gauge symmetry groups] may be the only fundamental mediators in the sense that they are not composite of other elementary entities. This is another reason to consider that these interactions may originate from a single fundamental interaction.

It is the purpose of this paper to consider such a unification. Although it is possible to unify QED with QCD, but the scale at which it happens is low enough to make proton life-time apparently inconsistent with the experimental limit. Other consequence of this unification is the mass relations between quarks and leptons which are in reasonable agreement with experimental values.

It is natural to take gauge group for the unification of QED with QCD as Pati-Salam group [1] SU_c (4). The fermions are then assigned to representations 4 and 6 of this group :

$$\psi_d = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ e^+ \end{pmatrix}, \quad \psi_u = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & u_3^c & -u_2^c & -u_1 \\ -u_3^c & 0 & u_1^c & -u_2 \\ u_2^c & -u_1^c & 0 & -u_3 \\ u_1 & u_2 & u_3 & 0 \end{pmatrix} \tag{1}$$

The superscript c denotes a charge conjugate state. The subscripts 1, 2, 3 denote the three coloured states. We note that

$$\bar{4} \times 4 = 1 + 15$$
$$4 \times 4 = 6_A + 10_S$$

$$\bar{6} \times 6 = 1 + 15 + 20''$$

There are 15 gauge vector bosons belonging to adjoint representation of SU_c (4). The charge operator for the fundamental representation 4 is given by

$$Q = \frac{2}{\sqrt{6}} \lambda_{15} = \frac{4}{\sqrt{6}} F_{15},$$

where $F_i = (1, \ldots, 15)$ are generators of the group SU_c (4). The co-variant derivative for the fundamental representation is given by

$$D_\mu = \partial_\mu + \frac{i}{2} g_G \underset{\sim}{\lambda} \cdot \underset{\sim}{G}_\mu$$

$$\frac{1}{2} \underset{\sim}{\lambda} \cdot \underset{\sim}{G}_\mu = G_\mu \tag{3}$$

G_μ is a 4×4 matrix representing 15 gauge vector bosons. It then follows that gauge invariant interaction Lagrangian is given by

$$L_{int} = i \frac{g_G}{\sqrt{2}} \overline{\psi}_d \gamma_\mu G_\mu \psi_d$$

$$+ \frac{i g_G}{\sqrt{2}} \left(\mathrm{Tr} \left(\overline{\psi}_u \gamma_\mu G_\mu \psi_u \right) + \mathrm{Tr} \left(\overline{\psi}_u \gamma_\mu \psi_u G_\mu \right) \right) \tag{4}$$

The symmetry is spontaneously broken by introducing Higgs scalars H belonging to adjoint representation 15 of SU_c (4) :

$$SU_c(4) \rightarrow SU_c(3) \times U_{em}(1)$$

We can represent the 15-plet Higgs scalars as 4×4 traceless matrix H. Then its vacuum expectation value can be put in the diagonal form :

$$\langle H \rangle = v \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -3 \end{bmatrix} \tag{5}$$

The gauge invariant Lagrangian for Higgs scalar is given by

$$L^H_{int} = -\frac{1}{2} D_\mu H_i D_\mu H_i, \tag{6}$$

where

$$D_\mu H_i = \partial_\mu H_i + g_G f_{ijk} G_{j\mu} H_k$$

By breaking symmetry in this way, the 15 gauge vector bosons split into the following representation of SU_c (3).

$G_\mu \equiv 8$

$X_\mu \equiv 3, \overline{X}_\mu \equiv \overline{3}$

A_μ = photon

The eight coloured gluons G_μ and photon A_μ remain massless, but 6 lepto-quarks X_μ's have acquired superheavy masses. From Eq. (6), L_{mass} for gauge vector bosons is given by

$$L_{mass} = \frac{1}{2} g_G^2 [\mathrm{Tr}\,(G_\mu H G_\mu H) - \mathrm{Tr}\,(HH G_\mu G_\mu)]$$

$$= -\frac{1}{4} \cdot g_G^2 (16 v^2)\, 2\, [\overline{X}_{1\mu} X_{1\mu} + \overline{X}_{2\mu} X_{2\mu} + \overline{X}_{3\mu} X_{3\mu}]$$

Hence we see that six lepto-quarks X_μ's have become superheavy, with a common mass,

$$M_X^2 = 8\, g_G^2\, v^2 \tag{8}$$

whereas 8 coloured gluons and photon remain massless.

In the broken symmetry, the interaction Lagrangian is given by

$$L_{int} = i\, g_s\, [\sum_{q=u,\, d} \overline{q}\, \gamma_\mu \frac{1}{2} \underset{\sim}{\lambda}\, q\,] \cdot \underset{\sim}{G}_\mu$$

$$+ i\, e\, [\sum_{i=1,\,2,\,3} (-\frac{1}{3} \overline{d}_i\, \gamma_\mu\, d_i + \frac{2}{3} \overline{u}_i \cdot \gamma_\mu\, u_i)$$

$$+ \overline{e^+}\, \gamma_\mu\, e^+\,]\, A_\mu$$

$$+ \frac{i\, g_G}{\sqrt{2}} [\sum_i \overline{d}_i\, \gamma_\mu\, e^+\, X_{i\mu} + \text{h.c.}]$$

$$- \frac{i\, g_G}{\sqrt{2}} [\sum_i \overline{e^-}\, \gamma_\mu\, d_i^c\, X_{i\mu} + \text{h.c.}]$$

$$- \frac{i\, g_G}{\sqrt{2}} [\,\varepsilon_{ijk}\, \overline{u_i^c}\, \gamma_\mu\, u_j\, X_{k\mu} + \text{h.c.}] \tag{9}$$

In the symmetry limit, viz. at unifying mass M_X

$$g_s^2 \, (M_X) = g_G^2 \tag{10}$$

$$e^2 \, (M_X) = \frac{3}{8} g_G^2 \tag{11}$$

The fermion mass term arises from two sources. Since, it is a vector theory, the "bare' mass term is also possible ; the other source of mass is the spontaneous symmetry breaking. For *d*-quarks and electron, the mass term is given by

$$-m_{01} \left[\sum_i \bar{d}_i \, d_i + \bar{e}e \right] - f_1 \, v \left[\sum_i \bar{d}_i \, d_i - 3\bar{e}e \right] \tag{12}$$

Hence, we have

$$m_d = m_{01} + f_1 \, v \tag{13}$$

$$m_e = m_{01} - 3 \, f_1 \, v$$

For *u*-quarks, the bare mass term is given by

$$-m'_{01} \, \mathrm{Tr} \, (\overline{\psi}_u \, \psi_u)$$

$$= - \, m'_{01} \left[\sum_i (\bar{u}_i \, u_i + \bar{u}^c_i \, u^c_i) \right] \tag{14}$$

The mass of *u*-quark could also arise from an invariant of the form $\varepsilon^{\alpha\beta\gamma\delta} \, \psi^T_{\gamma\delta} \, C \, \psi_{\alpha\beta}$, but it can be understood to be included in m'_{01}. From the 15-plet Higgs scalars, there is no contribution to mass term for *u*-quarks.

From Eq. (14), we have then

$$m_u = m'_{01} \tag{15}$$

We have written all our equations for the first generation (*u*, *d*,*e*). But these equations are of course applicable for other generations (*c*, *s*) and (*t*, *b*). The mass relations given in Eqs. (13) and (15) are at the unification mass scale M_X and are subject to renormalization.

We now derive the scale at which this unification occurs. We note that fine structure constant is given by [2]

$$\alpha^{-1}(M_X) = \alpha^{-1} - \frac{2}{3\pi}\sum_f Q_f^2 \ln \frac{M_X}{m_f} \tag{16}$$

For the three generations, with 3 charged leptons, three coloured triplets of quarks with $(-1/3)$ unit of charge and three coloured triplets of quarks with (2/3) units of charge, $\sum_f Q_f^2 = 8$. For $SU_c(3)$, we have

$$\alpha_s(M_X) = \frac{12\pi}{2\times(33-2N_f)\ln M_X/\Lambda} = \frac{4\pi}{7\times 2\ln M_X/\Lambda} \tag{17}$$

where Λ is QCD constant of few hundred MeV ($\Lambda \sim 0.2$ GeV to 0.5 GeV).

Now using Eqs. (16), (17) and (11), we have

$$\ln \frac{M_X}{\Lambda} = \frac{3\pi}{44\alpha}\left(1 - \frac{2\alpha}{3\pi}\sum_f Q_f^2 \ln \frac{\Lambda}{m_f}\right) \tag{18}$$

From this equation, using $\Lambda = 0.5$ GeV, we have

$$M_X = 3.5 \times 10^{12} \text{ GeV approximately} \tag{19}$$

Similar result follows from the renormalization equations considered below. It is convenient to define β_3 $[N_f = 6]$ and β_{em} $[\sum Q_f^2 = 8]$, where

$$\beta_3 = \frac{1}{4\pi}\left(-11 + \frac{2}{3}N_f\right) = -\frac{7}{4\pi}$$

$$\beta_{em} = \frac{1}{4\pi}\left(\frac{32}{3}\right) = \frac{32}{12\pi} \tag{20}$$

Then, we have

$$\alpha_s^{-1}(\mu) = \alpha_G^{-1} + 2\beta_3 \ln \frac{M_X}{\mu}$$

$$\alpha^{-1}(\mu) = \frac{8}{3}\alpha_G^{-1} + 2\beta_{em} \ln \frac{M_X}{\mu} \tag{21}$$

From Eqs. (20) and (21), we have

$$\ln \frac{M_X}{\mu} = \frac{3\pi}{44\alpha}\left(1 - \frac{8}{3}\frac{\alpha}{\alpha_s(\mu)}\right) \tag{22}$$

With $\alpha_s(\mu) = 0.2$, $\mu = 10$ GeV, we have again

$$M_X = 3.2 \times 10^{12} \text{ GeV} \tag{23}$$

The masses of fermions are also subject to renormalization. The masses given in Eqs. (13) and (15) are the masses at M_X. The masses at $\mu = 10$ GeV (say) are given by [3] (with three generations).

$$\frac{m_q(\mu)}{m_q(M_X)} = \left(\frac{\alpha_s(\mu)}{\alpha_G}\right)^{4/7}\left(\frac{8\alpha(\mu)}{3\alpha_G}\right)^{-1/32}, \quad q = d, s, b \tag{24}$$

$$\frac{m_l(\mu)}{m_l(M_X)} = \left(\frac{8}{3}\frac{\alpha(\mu)}{\alpha_G}\right)^{-9/32}, \qquad l = e, \mu, \tau \tag{25}$$

$$\frac{m'_q(\mu)}{m'_q(M_X)} = \left(\frac{\alpha_s(\mu)}{\alpha_G}\right)^{4/7}\left(\frac{8}{3}\frac{\alpha(\mu)}{\alpha_G}\right)^{-4/32}, \quad q' = u, c, t \tag{26}$$

From Eqs. (24) and (25),

$$\frac{m_q(\mu)}{m_l(\mu)} = \frac{m_q(M_X)}{m_l(M_X)}\left(\frac{\alpha_s(\mu)}{\alpha_G}\right)^{4/7}\left(\frac{8}{3}\frac{\alpha(\mu)}{\alpha_G}\right)^{8/32},$$

$$\equiv a\, m_q(M_X) / m_l(M_X) \tag{27}$$

where

$$a = (\alpha_s(\mu)/\alpha_G)^{4/7}\left(\frac{8}{3}\frac{\alpha(\mu)}{\alpha_G}\right)^{8/32} \tag{28}$$

For $\alpha_s(\mu) = 0.2$, $M_X \sim 3.2 \times 10^{12}$ GeV approximately, $\alpha_G \sim 1/34.5$, $a = 2.7$. The finite mass effects tend to reduce the quantity a. It is therefore good estimate to take $a \sim 2.5$.

There are three layers of matter, the bottom layer of the matter being the heaviest. It is, therefore, natural to take

$$m_{o3} >> m_{o2} > m_{01} \tag{29}$$

We will now assume that Higgs scalars are coupled to each generation of fermions with a universal coupling so that $f_1 = f_2 = f_3 \equiv f$.

Then we have

$$\frac{fv}{m_{a3}} << \frac{fv}{m_{a2}} < \frac{fv}{m_{a1}} = r \tag{30}$$

Let us define,

$$R_{31} = \frac{m_{a3}}{m_{c1}}$$

$$R_{21} = \frac{m_{a2}}{m_{a1}} \tag{31}$$

We now estimate r, R_{31} and R_{21}. For this purpose we take $\frac{m_d}{m_s} \sim \frac{1}{20}$ and use experimental ratios for $\frac{m_e}{m_\mu}$ and $\frac{m_e}{m_\tau}$. Then from Eqs. (24) and (25),

$$\frac{m_d}{m_s} = \frac{m_d\,(M_X)}{m_s\,(M_X)} \sim 1/20$$

$$\frac{m_e}{m_\mu} = \frac{m_e\,(M_X)}{m_\mu\,(M_X)} \sim 1/200$$

$$\frac{m_e}{m_\tau} = \frac{m_e\,(M_X)}{m_\tau\,(M_X)} \sim 2.8 \times 10^{-4} \tag{32}$$

Thus from Eqs. (32) and (12), one obtains

$$r \sim 0.29$$

$$R_{21} \sim 25.5$$

$$R_{31} \sim 460 \tag{33}$$

Hence from Eqs. (27), we have

$$\frac{m_d}{m_e} = a\,\frac{1+r}{1-3r} \sim 9.9a$$

$$\frac{m_s}{m_\mu} = a\,\frac{1 + r/R_{21}}{1 - 3r/R_{21}} \sim 1.05a$$

$$\frac{m_b}{m_\tau} = a\,\frac{1 + r/R_{31}}{1 - 3r/R_{31}} \sim a \qquad (34)$$

Then the quark masses come out to be as follows with $a = 2.5$

$$m_d = 12.4 \text{ MeV}$$

$$m_s = 276 \text{ MeV}$$

$$m_b = 4.5 \text{ GeV} \qquad (35)$$

These mass relations are in good agreement with their experimental values.

The lepto-quarks X's carry an electric charge of 4/3 units. They give rise to processes for which $\triangle B = 1$, $\triangle L = 1$. The effective four fermion interaction for such processes is given by (see Eq. (9))

$$\frac{g_G^2}{2M_X^2}\,[\,(\varepsilon_{ijk}\,\bar{u}_i^c\,\gamma_\mu\,u_j)\,(\bar{d}_k\,\gamma_\mu\,e^+\,) + h.c.\,] \qquad (36)$$

Such an interaction is responsible for proton decay of the type

$$p \rightarrow e^+ + x^o$$

Since in this case $M_X \times 3\,2 \sim 10^{12}$ GeV, Eq. (36) gives a decay rate too fast to be consistent with the experimental limit on the proton life-time. However, proton decay may be avoided by suitable Cabbibo like rotations in lepto-quark couplings. This point will be considered elsewhere.

ACKNOWLEDGEMENT

One of us (R) would like to acknowledge the support of the Research Committee of the University of Petroleum and Minerals for this work.

REFERENCES

1. The group SU_c (4) (where the fourth colour is identified with lepton members) for unification of electroweak and strong interactions was first used by Pati and Salam (J.C. Pati and A. Salam, Phys. Rev. **D 10,** 275 (1974)). But our unification scheme has a resemblance to SU (5) grand unification model of Georgi and Glashow (H. Georgi and S.L. Glashow, Phys. Rev. Lett. **32**, 438 (1974)).

2. W.J. Marciano, Phys. Rev. **D 20**, 274 (1979), T.J. Goldman and D.A. Ross, Nucl. Phys. **B162,** 102 (1980).

3. A.J. Buras, J. Ellis, M.K. Gaillard and D.V. Nanopoulos, Nucl. Phys. **B135,** 660 (1978).

The Punjab University
Journal of Mathematics
Vol. XIV—XV (1981-82), pp. 51-70.

SPECTRAL REPRESENTATION OF THE LOVE WAVE OPERATOR FOR TWO LAYERS OVER A HALF-SPACE

By

M.H. KAZI & A.S.M. ABU-SAFIYA

Department of Mathematical Sciences
University of Petroleum and Minerals
Dhahran, Saudi Arabia

Summary.

In this paper we use the method described in Kazi [1976] to determine the spectral representation of the two-dimensional Love wave operator associated with the propagation of monochromatic SH waves in a structure consisting of two uniform layers overlying a uniform half-space. The representation is useful in tackling a class of Love wave diffraction problems in horizontally discontinuous structures involving three-layered models.

1. Introduction.

In a series of papers Kazi [1978a, b], Kazi [1979], Niazy and Kazi [1980], the authors use a method, based upon an integral equation formulation together with the application of Schwinger-Levine variational principle to investigate the two-dimensional problems of the propagation of plane, harmonic, monochromatic Love waves, incident normally upon the plane of discontinuity in laterally discontinuous structures involving step-wise change in surface topography or change in material properties. Diffraction of Love waves is described by means of a scattering matrix and approximate expressions for its elements are sought through the variational principle. Reflection and transmission coefficients are then obtained through a transmission matrix related to the scattering matrix. The method has the advantage that it takes into account the body-wave contributions. However, the method pre-supposes the existence of a complete set of proper or improper eigenfunctions, in terms of which the displacements on either side of the discontinuity may be expressed. In order to accomplish this Kazi [1976] gave a general method for finding the spectral reyresentation of the two-dimensional Love wave operator associated with the propagation of monochromatic SH waves in a laterally-uniform layered strip or half-space. However, specific spectral representations were found only for two-layer models. Recently, Kennett [1981] has discussed the spectral representation of the elastodynamic operator associated with coupled seismic waves.

In this paper, we follow the same procedure as in Kazi [1976] to determine the explicit spectral representation of the Love wave operator associated with monochromatic SH-waves for a three-layer model comprising two homogeneous, infinite strips overlying a uniform half-space. This representation will find usage in tackling a class of Love wave diffraction problems associated with three-layer models by the method described above.

2. **Equations of motion.**

We wish to represent the two-dimensional motion of a laterally homogeneous structure consisting of two uniform layers over a uniform half-space in a general way ; the motion will consist of waves propagating along the direction of the x-axis in the coordinate system shown in Fig. 1. We consider a layer of infinite depth, rigidity μ_3, shear velocity β_3 and density ρ_3, overlaid by two infinite strips, consisting of a layer of finite depth H_2, density ρ_2, rigidity μ_2 ($< \mu_3$) and shear velocity β_2 ($< \beta_3$), and another layer of depth H_1 ($< H_2$), density ρ_1, rigidity μ_1 ($< \mu_2$) and shear velocity β_1 ($< \beta_2$) (see Fig. 1). We suppose the density and the rigidity of each layer to be constant, and the top plane surface to be stress-free.

We choose the axes in such a way that the upper free surface coincides with the plane $z = -H_1$ and the xy-plane coincides with the plane of welded contact between the two upper layers as shown in Fig. 1.

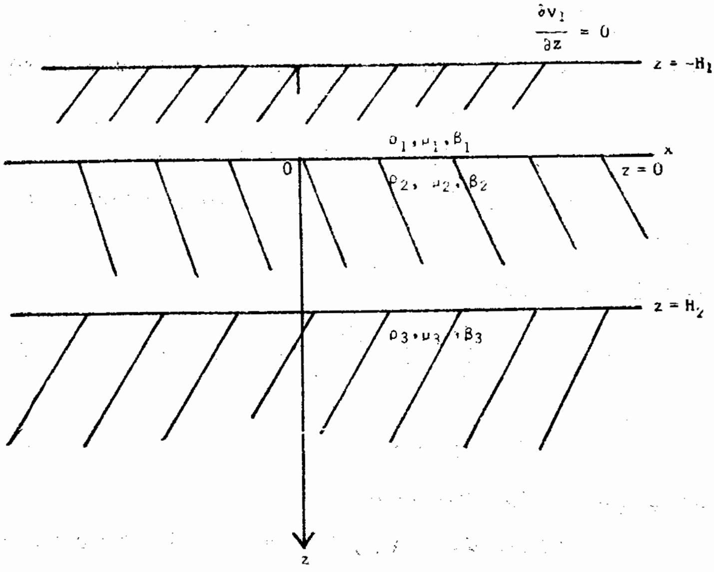


Figure 1. The geometry of the problem

We shall consider horizontally polarized shear waves only, which means that there are no displacements in the x and z directions and the motion is in the y-direction only. Let $v\ (x, z, t)$ be the y-component of displacement. It must satisfy the differential equation.

$$\rho\ (z)\ \frac{\partial^2 v}{\partial t^2} = \frac{\partial}{\partial x}\left(\mu\ (z) \frac{\partial v}{\partial x} \right) + \frac{\partial}{\partial z}\left(\mu\ (z)\ \frac{\partial v}{\partial z} \right), \tag{2.1}$$

where

$$\begin{aligned} \mu\ (z) &= \mu_1, \ -H_1 < z < 0 \\ &= \mu_2, \ 0 < z < H_2 \\ &= \mu_3, \ H_2 < z \end{aligned} \tag{2.2}$$

and

$$\begin{aligned} \rho\ (z) &= \rho_1, \ -H_1 < z < 0 \\ &= \rho_2, \ 0 < z < H_2 \\ &= \rho_3, \ H_2 < z \end{aligned} \tag{2.3}$$

For convenience we label the intervals $\{ z : -H_1 < z < 0 \}$, $\{z : 0 < z < H_2\}$ and $\{z : z > H_2\}$ as I_1, I_2 and I_3 respectively.

In order to obtain a general representation, we first of all examine harmonic waves, travelling in x-direction with positive real frequency ω and wave numbers k :

$$v(x, z, t) = V(z) \exp [i(\omega t - kx)]. \tag{2.4}$$

(We shall assume ω to be fixed and choose k to satisfy the propagation conditions).

Equation (2.1) becomes

$$L(V) = \frac{d}{dz}\left(\mu(z) \frac{dV}{dz} \right) + (\omega^2 \rho(z) - k^2 \mu(z))\, V = 0. \tag{2.5}$$

$$V(z) = V_i(z), \quad z \, \varepsilon \, I_i, \quad i = 1, 2, 3,$$

L being the *Love wave operator.*

$V_1(z)$, $V_2(z)$ and $V_3(z)$ satisfy the following equations :

$$\frac{d^2V_1}{dz^2} + \sigma_1^2 V_1 = 0, \quad \sigma_1^2 = \left(\frac{\omega^2}{\beta_1^2} - \lambda \right), \; \lambda = k^2,$$

$$\beta_1^2 = \frac{\mu_1}{\rho_1}, \; -H_1 < z < 0 \tag{2.6}$$

$$\frac{d^2V_2}{dz^2} + \sigma_2^2 V_2 = 0, \; \sigma_2^2 = \left(\frac{\omega^2}{\beta_2^2} - \lambda \right),$$

$$\beta_2^2 = \frac{\mu_2}{\rho_2}, \; 0 < z < H_2 \tag{2.7}$$

$$\frac{d^2V_3}{dz^2} - \sigma_3^2 V_3 = 0, \quad \sigma_3^2 = \left(\lambda - \frac{\omega^2}{\beta_3^2} \right), \; \beta_3^2 = \frac{\mu_3}{\rho_3}, \; H_2 < z \tag{2.8}$$

with the *interface conditions* :

$$V_1(0) = V_2(0) \tag{2.9}$$

$$\mu_1 V_1'(0) = \mu_2 V_2'(0) \tag{2.10}$$

and

$$V_2(H_2) = V_3(H_2) \tag{2.11}$$

$$\mu_2 V_2'(H_2) = \mu_3 V_3'(H_2) \tag{2.12}$$

and the *boundary conditions.*

$$V_1'(-H_1) = 0,$$

$$\int_{-H_1}^{\infty} \mu(z) \mid V(z) \mid^2 dz < \infty, \tag{2.14}$$

where $\mu(z)$ is given by (2.2) and (′) denotes differentiation with respect to z.

The system (2.5), (2.9) – (2.14) is a SINGULAR Sturm-Liouville system with two points of discontinuity and corresponding interface conditions. Such systems have been discussed in detail in Kazi [1976] The boundary condition at infinity (2.14) is taken to be the requirement that the solution must be of finite μ-norm, so as to ensure the uniqueness of the so'ution as explained in Kazi [1976].

3. Green's function

Let $G(z, \zeta; \lambda) \mid_{z \in I_i,\, \zeta \in I_j} = G_{ij}$, where $i, j = 1, 2, 3$ (see Fig. 2). G_{ij} determine the Green function $G(z, \zeta; \lambda)$ completely, provided the following conditions are satisfied :

(G_1) $G_{ij}(z, \zeta; \lambda)$ is a continuous function of z for all $z \in I_i$.

(G_2) $G_{ij}(z, \zeta; \lambda)$ $(i \neq j)$ possesses a continuous first order derivative of z at each point of I_i ; G_{ij} $(i = j)$ possesses a continuous first order derivative at each point of I_i except $z = \zeta$, where it has a jump discontinuity, given by :

$$G'_{ii}(\zeta^+, \zeta; \lambda) - G'_{ii}(\zeta^-, \zeta; \lambda) = \frac{1}{\mu_i(\zeta)}.$$

(G_3) If $i \neq j$, $L(G_{ij}) = 0$. If $i = j$, $L(G_{ij}) = 0$ for $z \neq \zeta$.

(G_4) $G(z, \zeta; \lambda)$ satisfy the *interface conditions* (2.9) – (2.12) and the *boundary conditions* (2.13) – (2.14).

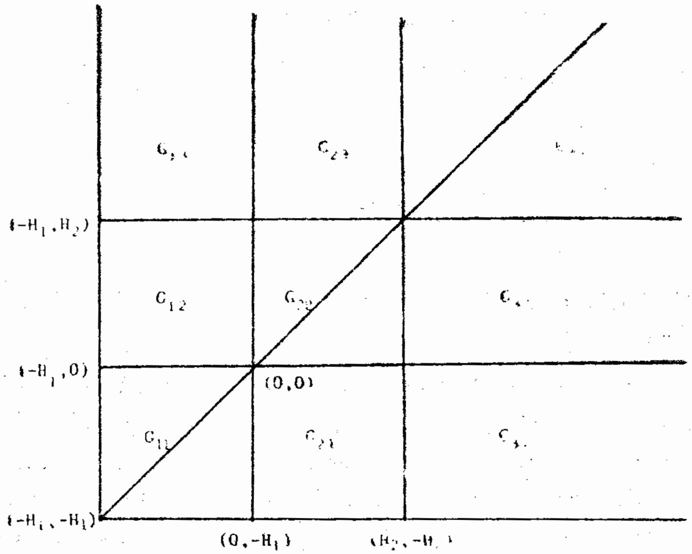


Figure 2 The character of the Green function

The Green function is unique and symmetric. We now proceed to construct the Green function explicitly.

(*i*) If $\zeta \varepsilon I_1$, then G_{11}, G_{21} and G_{31} satisfy the differential equations :

$$\frac{\partial^2 G_{11}}{\partial z^2} + \sigma_1^2 G_{11} = \delta(z - \zeta) \tag{3.1}$$

$$\frac{\partial^2 G_{21}}{\partial z^2} + \sigma_2^2 G_{21} = 0 \tag{3.2}$$

and

$$\frac{\partial^2 G_{31}}{\partial z^2} - \sigma_3^2 G_{31} = 0 \tag{3.3}$$

together with the following conditions

$$G_{11}' = 0 \text{ at } z = -H_1 \tag{3.4a}$$

$$G_{11} = G_{21} \text{ at } z = 0 \tag{3.4b}$$

$$\mu_1 G_{11}' = \mu_2 G_{21}' \text{ at } z = 0 \tag{3.4c}$$

$$G_{21} = G_{31} \text{ at } z = H_2 \tag{3.4d}$$

$$\mu_2 G'_{21} = \mu_3 G'_{31} \quad \text{at } z = H_2 \tag{3.4e}$$

$$G_{11}(\zeta^+, \zeta; \lambda) = G_{11}(\zeta^-, \zeta; \lambda) \tag{3.4f}$$

$$G'_{11}(\zeta^+, \zeta; \lambda) - G'_{11}(\zeta^-, \zeta; \lambda) = \frac{1}{\mu_1} \tag{3.4g}$$

and

$$\int_{-H_1}^{\infty} \mu(z) \mid G(z) \mid^2 dz < \infty. \tag{3.4h}$$

After considerable effort, we find

$$G_{11} = \frac{\mu_2 \sigma_2 \cos \sigma_1 (\zeta + H_1) \cos \sigma_1 (z + H_1)}{\triangle \cos^2 \sigma_1 H_1} \left(1 + \frac{\mu_3 \sigma_3}{\mu_2 \sigma_2} \tan \sigma_2 H_2\right)$$
$$+ \frac{1}{\mu_1 \sigma_1 \cos \sigma_1 H_1} \{\cos \sigma_1 (z + H_1) \sin (\sigma_1 \zeta) \, \theta (\zeta - z)$$
$$+ \cos \sigma_1 (\zeta + H_1) \sin (\sigma_1 z). \, \theta (z - \zeta)\}, \; z \, \varepsilon \, I_1 \tag{3.5}$$

where

$$\triangle = \mu_1 \sigma_1 \mu_2 \sigma_2 \tan (\sigma_1 H_1) + \mu_1 \sigma_1 \mu_3 \sigma_3 \tan (\sigma_2 H_2) \times$$
$$\tan (\sigma_1 H_1) \quad - \mu_3 \sigma_3 \mu_2 \sigma_2 + \mu_2^2 \sigma_2^2 \tan (\sigma_2 H_2) \tag{3.6}$$

and

$$\theta (\zeta - z) = 1, \; \zeta > z$$
$$= 0, \; \zeta < z$$

is the Heaviside unit function ;

$$G_{21} = \frac{\cos \sigma_1 (\zeta + H_1)}{\triangle \cos (\sigma_2 H_2) \cos (\sigma_1 H_1)} \{\mu_2 \sigma_2 \cos (\sigma_2 (z - H_2))$$
$$- \mu_3 \sigma_3 \sin (\sigma_2 (z - H_2))\}, \; z \, \varepsilon \, I_2, \tag{3.7}$$

and

$$G_{31} = \frac{\mu_2 \sigma_2 \cos (\sigma_1 (\zeta + H_1))}{\triangle \cos (\sigma_2 H_2) \cos (\sigma_1 H_1)} \, e^{-\sigma_3 (z - H_2)},$$
$$z \, \varepsilon \, I_3. \tag{3.8}$$

(*ii*) If $\zeta \varepsilon I_2$, then G_{12}, G_{22} and G_{32} satisfy the differential equations

$$\frac{\partial^2 G_{12}}{\partial z^2} + \sigma_1^2 G_{12} = 0 \tag{3.9}$$

$$\frac{\partial^2 G_{22}}{\partial z^2} + \sigma_2^2 G_{22} = \delta (z - \zeta) \tag{3.10}$$

and

$$\frac{\partial^2 G_{32}}{\partial z^2} - \sigma_3^2 G_{32} = 0 \tag{3.11}$$

together with the conditions (3.4a) – (3.4h) suitably modified. We obtain

$$G_{12} = \frac{\cos \{\sigma_1 (z + H_1)\} . B(\zeta)}{\triangle \cos (\sigma_1 H_1) \cos (\sigma_2 H_2)}, \quad z \varepsilon I_1 \tag{3.12}$$

where $\triangle$ is given by (3.6) and

$$B(\zeta) = \mu_2 \sigma_2 \cos (\sigma_2 (\zeta - H_2)) - \mu_3 \sigma_3 \sin (\sigma_2 (\zeta - H_2)), \tag{3.13}$$

$$G_{22} = \frac{B(\zeta) \, B(z)}{M \triangle \cos^2 (\sigma_2 H_2)}$$

$$- \left\{ \frac{B(\zeta) \sin (\sigma_2 z) . \theta (\zeta - z) + B(z) . \sin (\sigma_2 \zeta) . \theta (z - \zeta)}{M \mu_2 \sigma_2 \cos (\sigma_2 H_2)} \right\}, \; z \varepsilon I_2 \tag{3.14}$$

where $\triangle$, $B(z)$ are given by (3.6), (3.13) and

$$M = \mu_2 \sigma_2 + \mu_3 \sigma_3 \tan \sigma_2 H_2,$$

$$G_{32} = \frac{1}{\triangle \cos \sigma_2 H_2} \times$$

$$\{\mu_2 \sigma_2 \cos (\sigma_2 \zeta) - \mu_1 \sigma_1 \tan (\sigma_1 H_1) \sin (\sigma_2 \zeta)\} \, e^{-\sigma_3 (z - H_2)},$$

$$z \varepsilon I_3 \tag{3.16}$$

(*iii*) If $\zeta \varepsilon I_3$, then G_{13}, G_{23} and G_{33} satisfy the differential equations

$$\frac{\partial^2 G_{13}}{\partial z^2} + \sigma_1^2 G_{13} = 0 \tag{3.17}$$

$$\frac{\partial^2 G_{23}}{\partial z^2} + \sigma_1^2 G_{23} = 0 \tag{3.18}$$

and

$$\frac{\partial^2 G_{33}}{\partial z^2} - \sigma_3^2 G_{33} = \delta(z-\zeta) \tag{3.19}$$

together with the conditions (3.4a) – (3.4h) suitably modified. We obtain

$$G_{13} = \frac{\mu_2 \sigma_2 \cos\{\sigma_1 (z+H_1)\}}{\triangle \cos(\sigma_1 H_1) \cos(\sigma_2 H_2)} e^{-\sigma_3(\zeta - H_2)}, \quad z \,\varepsilon\, I_1, \tag{3.20}$$

$$G_{23} = \frac{B(z)\, \mu_2 \sigma_2 e^{-\sigma_2(\zeta - H_2)}}{M \triangle \cos^2(\sigma_2 H_2)} - \frac{e^{-\sigma_3(\zeta - H_2)}}{M \cos \sigma_2 H_2}, \quad z \,\varepsilon\, I_2 \tag{3.21}$$

and

$$G_{33} = \frac{\mu_2^2 \sigma_2^2 e^{-\sigma_3(\zeta - H_2)} e^{-\sigma_3(z - H_2)}}{M \triangle \cos^2(\sigma_2 H_2)}$$

$$- \frac{e^{-\sigma_3(\zeta - H_2)} e^{-\sigma_3(z - H_2)} \{\mu_3 \sigma_3 \tan(\sigma_2 H_2) - \mu_2 \sigma_2\}}{2\mu_3 \sigma_3 M}$$

$$- \left\{ \frac{e^{-\sigma_3(\zeta - z)}}{2\mu_3 \sigma_3} \cdot \theta(\zeta - z) + \frac{e^{-\sigma_3(z-\zeta)}}{2\mu_3 \sigma_3} \cdot \theta(z-\zeta) \right\}$$

$$z \,\varepsilon\, I_3, \tag{3.22}$$

where $\triangle$, $B(z)$ and M are given by (3.6), (3.13) and (3.15) respectively.

We note that $G_{ij}(z, \zeta\,; \lambda) = G_{ji}(\zeta, z\,; \lambda)$, $i, j = 1, 2, 3$ i.e. the Green function is symmetric.

4. Spectral representation :

The essential step in obtaining the spectral representation is to integrate the Green function $G(z, \zeta\,; \lambda)$ obtained in the previous section around a large circle $|\lambda| = R$ in the complex λ-plane. The Green function has, in addition to simple poles, a branch-point singularity. The spectrum is the disjoint union of the point-spectrum, giving rise to proper eigenfunctions, and the continuous spectrum, which yields improper eigenfunctions. The continuous spectrum will be the set of points on the branch-cut along a portion of the real axis and the discrete spectrum will be the set of poles lying on the real axis. The sum of residues at the poles and the contribution from the branch-cut yield the following representation of the delta function in terms of proper eigenfunctions $\{\phi^{(n)}(z)\}$ and improper eigenfunctions $\{\psi(z,\lambda)\}$ (see Kazi [1976]) ; such a representation is useful, because it enables us to find the corresponding eigenfunction from the knowledge of Green's function.

$$\lim_{R\to\infty} \frac{-1}{2\pi i} \oint_{|\lambda|=R} G(z,\zeta;\lambda)\, d\lambda = \sum_n \phi^{(n)}(z)\, \overline{\phi^{(n)}(\zeta)} + \int \psi(z,\lambda)\, \overline{\psi(\zeta,\lambda)}\, d\lambda = \frac{\delta(z-\zeta)}{\mu(\zeta)} \qquad (4.1)$$

(*i*) First, we consider

$$I_{11} = \lim_{R\to\infty} \frac{-1}{2\pi i} \oint_{|\lambda|=R} G_{11}(z,\zeta;\lambda)\, d\lambda$$

$$= \lim_{R\to\infty} \frac{-1}{2\pi i} \oint_{|\lambda|=R} \Big[M \frac{\cos\{\sigma_1(\zeta+H_1)\}\cos\{\sigma_1(z+H_1)\}}{\Delta \cos^2\sigma_1 H_1}$$

$$+ \frac{1}{\mu_1\sigma_1\cos\sigma_1 H_1} \{\cos(\sigma_1(z+H_1))\sin(\sigma_1\zeta)\,\theta(\zeta-z)$$

$$+ \cos(\sigma_1(\zeta+H_1))\sin(\sigma_1 z)\,\theta(z-\zeta)\}\Big]\, d\lambda, \quad \text{(using 3.5)}, \qquad (4.2)$$

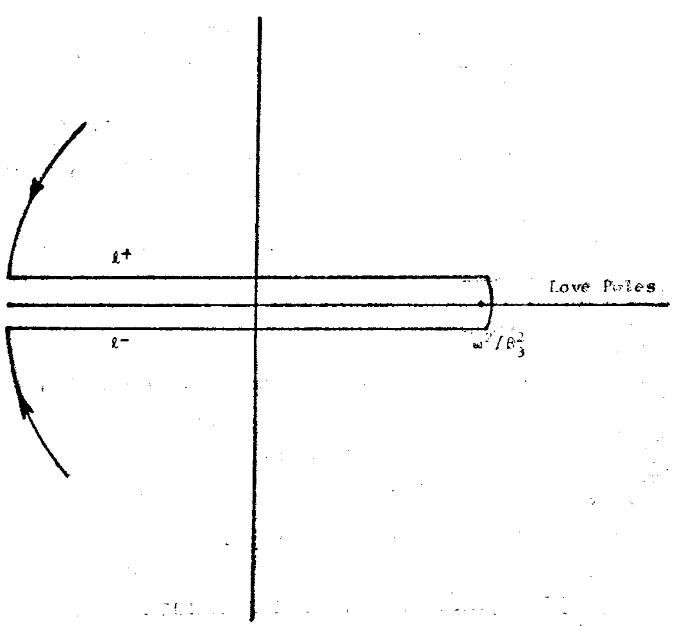


Figure 3. The contour of integration in the complex λ-plane.

where Δ and M are given by (3.6) and (3.15) respectively. We note that $\lambda = \omega^2/\beta_3^2$ is the only branch-point singularity of the integrand of (4.2) and the poles are the roots of

$$\Delta = \mu_1 \sigma_1 \mu_2 \sigma_2 \tan(\sigma_1 H_1) + \mu_1 \sigma_1 \mu_3 \sigma_3 \tan(\sigma_2 H_2) \tan(\sigma_1 H_1) - \mu_3 \sigma_3 \mu_2 \sigma_2 + \mu_2^2 \sigma_2^2 \tan(\sigma_2 H_2) = 0 \tag{4.3}$$

which is the *dispersion equation* for Love wave propagation in two layers over a half-space [see Ewing et al. 1957, p. 229]. The poles are all simple, finite in number and are located in the interval $(\omega^2/\beta_3^2, \omega^2/\beta_1^2]$. The continuous spectrum is related to the integral over the branch-lines l^+, l^- in the complex λ-plane, and the path of integration is shown in Fig. 3. These remarks are valid for all the integrands we shall encounter. We assume that Re $(\sigma_3) > 0$ for I $(\lambda) \neq 0$. This means that on the branch-line l^+, $\sigma_3 = is_3$ and on l^-, $\sigma_3 = -is_3$, where $s_3 = (\omega^2/\beta_3^2 - \lambda)^{\frac{1}{2}}$ is real and positive for $\lambda < \omega^2/\beta_3^2$.

Let

$$\gamma_1 = \frac{M}{\Delta} = \frac{\mu_2 \sigma_2 + \mu_3 \sigma_3 \tan(\sigma_2 H_2)}{\Delta}.$$

Then

$$\gamma_1^+ - \gamma_1^- = 2i \; I(\gamma_1^+) = \frac{2i\, \mu_2^2 \sigma_2^2 \mu_3 s_3 \sec^2(\sigma_2 H_2)}{p^2 + q^2} \tag{4.4}$$

where

$$p = \mu_1 \sigma_1 \mu_2 \sigma_2 \tan(\sigma_1 H_1) + \mu_2^2 \sigma_2^2 \tan(\sigma_2 H_2) \tag{4.5}$$

$$q = \mu_1 \sigma_1 \mu_3 s_3 \tan(\sigma_2 H_2) \tan(\sigma_1 H_1) - \mu_2 \sigma_2 \mu_3 s_3, \tag{4.6}$$

and the superscripts + and − refer to the values at the branches l^+ and l^- respectively.

The contribution to I_{11} from $l+$ and $l-$ is

$$-\frac{1}{2\pi i}\left(\int_{l+} G_{11}\, d\lambda - \int_{l-} G_{11}\, d\lambda\right) = \frac{-1}{2\pi i}\int_{-\infty}^{\omega^2/\beta_3^2} (G_{11}^{+} - G_{11}^{-})\, d\lambda$$

$$= -\frac{1}{2\pi i}\int_{-\infty}^{\omega^2/\beta_3^2} \frac{(\gamma_1^{+} - \gamma_1^{-}) \cos(\sigma_1(\zeta+H_1)) \cos(\sigma_1(z+H_1))}{\cos^2 \sigma_1 H_1}\, d\lambda$$

$$= -\frac{\mu_3 \mu_2^2}{\pi}\int_{-\infty}^{\omega^2/\beta_3^2} \frac{s_3 \sigma_2^2 \cos(\sigma_1(\zeta+H_1)) \cos(\sigma_1(z+H_1))}{(p^2+q^2)\cos^2(\sigma_2 H_2)\cos^2(\sigma_1 H_1)}\, d\lambda,$$

(using (4.4))

$$= -\int_{-\infty}^{\omega^2/\beta_3^2} \psi_1(z, \lambda)\, \psi_1(\zeta, \lambda)\, d\lambda, \tag{4.7}$$

where

$$\psi_1(z, \lambda) = \frac{\mu_2 \sigma_2 \mu_3 s_3 \cos(\sigma_1(z+H_1)) \cos\theta}{\rho\sqrt{\pi \mu_3 s_3}\cos(\sigma_1 H_1)\cos(\sigma_2 H_2)} \tag{4.8}$$

and

$$\theta = \tan^{-1}(q/p,) \tag{4.9}$$

where p and q are given by (4.5) and (4.6) respectively. The sum of the residues at the poles $\{\lambda_n\}$ is given by

$$-\sum_{n=1}^{N} \frac{\cos(\sigma_1^{(n)}(\zeta+H_1))\cos(\sigma_1^{(n)}(z+H_1))(M)_{\lambda=\lambda_n}}{\cos^2(\sigma_1^{(n)} H_1)\left[\frac{\partial\Delta}{\partial\lambda}\right]_{\lambda=\lambda_n}}$$

$$= \sum_{n=1}^{N} \phi_1^{(n)}(z)\, \phi_1^{(n)}(\zeta) \tag{4.10}$$

where

$$\sigma_1^{(n)} = \left(\frac{\omega^2}{\beta_1^2} - \lambda_n \right)^{\frac{1}{2}}, \; \sigma_2^{(n)} = \left(\frac{\omega^2}{\beta_2^2} - \lambda_n \right)^{\frac{1}{2}},$$

$$\sigma_3^{(n)} = \left(\lambda_n - \frac{\omega^2}{\beta_3^2} \right)^{\frac{1}{2}} \quad \text{and}$$

$$\phi_1^{(n)}(z) = \frac{\cos \sigma_1^{(n)} (z + H_1)}{\cos (\sigma_1^{(n)} H_1)} \left[\left\{ \frac{M}{\frac{\partial}{\partial \lambda} (-\Delta)} \right\}_{\lambda=\lambda_n} \right]^{\frac{1}{2}} \tag{4.11}$$

From (4.2), (4.7) and (4.10) we obtain

$$I_{11} = \sum_{n=1}^{N} \phi_1^{(n)}(z) \, \phi_1^{(n)}(\zeta) - \int_{-\infty}^{\omega^2/\beta_2^2} \psi_1(z, \lambda) \, \psi_1(\zeta, \lambda) \, d\lambda, \tag{4,12}$$

where ψ_1 is given by (4.8) and $\phi_1^{(n)}$ is given by (4.11).

(*ii*) Next, we consider

$$I_{21} = \lim_{R \to \infty} \frac{-1}{2\pi i} \oint_{|\lambda| = R} G_{21}(z, \zeta; \lambda) \, d\lambda$$

$$= \lim_{R \to \infty} \frac{-1}{2\pi i} \oint_{|\lambda| = R} \frac{B(z) \cos \sigma_1 (\zeta + H_1)}{\Delta \cos (\sigma_2 H_2) \cos (\sigma_1 H_1)} \, d\lambda,$$

(using (3.7)) $\qquad$ (4.13)

where $B(z)$ is given by (3.13).

Let $\gamma_2 = \dfrac{B(z)}{\Delta}$. Then

$$\gamma_2^+ - \gamma_2^- = 2i \, I\left(\gamma_2^+ \right)$$

$$= \frac{2i \, \mu_2 \, \sigma_2 \, \mu_3 \, s_3 \, C(z)}{(p^2 + q^2) \cos (\sigma_2 H_2)}, \tag{4.14}$$

where

$$C(z) = \mu_2 \sigma_2 \cos(\sigma_2 z) - \mu_1 \sigma_1 \sin(\sigma_2 z) \tan(\sigma_1 H_1) \quad (4.15)$$

and p and q are given by (4.5) and (4.6) respectively.

The branch-line contribution to the integral is given by

$$-\frac{\mu_2 \mu_3}{\pi} \int_{-\infty}^{\omega^2/\beta_3^2} \frac{\sigma_2 s_3 \cos(\sigma_1 (\zeta + H_1)) C(z)}{(p^2 + q^2) \cos^2(\sigma_2 H_2) \cos(\sigma_1 H_1)} d\lambda$$

$$= -\int_{-\infty}^{\omega^2/\beta_3^2} \psi_2(z, \lambda) \psi_1(\zeta, \lambda) \, d\lambda \quad (4.16)$$

where $\psi_1(\zeta, \lambda)$ is given by (4.8) and

$$\psi_2(z, \lambda) = \frac{\mu_3 s_3 C(z) \cos\theta}{p \cos(\sigma_2 H_2) \sqrt{\pi \mu_3 s_3}} \quad (4.17)$$

where θ is given by (4.9) and $C(z)$ is given by (4.15).

Contribution from the poles is given by

$$-\sum_{n=1}^{N} \frac{\cos\{\sigma_1^{(n)} (\zeta + H_1)\} [B(z)]_{\lambda=\lambda_n}}{\cos(\sigma_2^{(n)} H_2) \cos(\sigma_1^{(n)} H_1) \left[\frac{\partial}{\partial\lambda} \Delta\right]_{\lambda=\lambda_n}}$$

$$= \sum_{n=1}^{N} \phi_2^{(n)}(z) \, \phi_1^{(n)}(\zeta), \quad (4.18)$$

where $\phi_1(\zeta)$ is given by (4.11) and

$$\phi_2^{(n)}(z) = \frac{1}{\cos(\sigma_2^{(n)} H_2)} \left[\left\{ \frac{1}{M \frac{\partial}{\partial\lambda}(-\Delta)} \right\}_{\lambda=\lambda_n} \right]^{\frac{1}{2}} \times$$

$$\{B(z)\}_{\lambda=\lambda_n} \quad (4.19)$$

From (4.13), (4.16) and (4.18) we obtain

$$I_{21} = \sum_{n=1}^{N} \phi_2^{(n)}(z)\, \phi_1^{(n)}(\zeta)$$

$$- \int_{-\infty}^{\omega^2/\beta_3^2} \psi_2(z, \lambda)\, \psi_1(\zeta, \lambda)\, d\lambda, \qquad (4.20)$$

where $\phi_1^{(n)}$, $\phi_2^{(n)}$, ψ_1 and ψ_2 are given by (4.11), (4.19), (4.8) and (4.17) respectively.

(*iii*) Next, we consider

$$I_{31} = \lim_{R\to\infty} \frac{-1}{2\pi i} \oint_{|\lambda|=R} G_{31}(z, \zeta\,;\lambda)\, d\lambda$$

$$= \lim_{R\to\infty} \frac{-1}{2\pi i} \oint_{|\lambda|=R} \frac{\mu_2\sigma_2 \cos(\sigma_1(\zeta+H_1))\, e^{-\sigma_3(z-H_2)}}{\Delta \cos(\sigma_2 H_2)\cos(\sigma_1 H_1)}\, d\lambda$$

(4.21)

Let

$$\gamma_3 = \frac{e^{-\sigma_3(z-H_2)}}{\Delta}.$$

Then

$$\gamma_3^+ - \gamma_3^- = 2i\, I\left(\gamma_3^+\right)$$

$$= \frac{-2i\, D(z)}{p^2 + q^2},$$

where

$$D(z) = p \sin\{s_3(z-H_2)\} + q\cos\{s_3(z-H_2)\} \qquad (4.22)$$

and p, q are given by (4.5) and (4.6) respectively.

The branch-line contribution is given by

$$\frac{\mu_2}{\pi}\int_{-\infty}^{\omega^2/\beta_3^2} \frac{\sigma_2 \cos \sigma_1 (\zeta + H_1) \cos \theta \,.\, \sin \{\theta + s_3 (z - H_2)\}}{p \cos (\sigma_2 H_2) \cos (\sigma_1 H_1)} d\lambda$$

$$= - \int_{-\infty}^{\omega^2/\beta_3^2} \psi_3 (z, \lambda)\, \psi_1 (\zeta, \lambda)\, d\lambda, \tag{4.23}$$

where

$$\psi_3 (z, \lambda) = \frac{- \sin (\theta + s_3 (z - H_2) \}}{\sqrt{\pi \mu_3 s_3}} \tag{4.24}$$

and ψ_1 is given by (4.8).

Contribution from the poles is given by

$$- \sum_{n=1}^{N} \frac{\mu_2 \sigma_2^{(n)} \cos \{\sigma_1^{(n)} (\zeta + H_1)\}\, e^{-\sigma_3^{(n)} (z - H_2)}}{\cos (\sigma_2^{(n)} H_2) \cos (\sigma_1^{(n)} H_1) \left[\frac{\partial}{\partial \lambda} \triangle\right]_{\lambda = \lambda_n}}$$

$$= \sum_{n=1}^{N} \phi_3^{(n)} (z)\, \phi_1^{(n)} (\zeta) \tag{4.25}$$

where

$$\phi_3^{(n)} (z) = \frac{\mu_2 \sigma_2^{(n)} e^{-\sigma_3^{(n)} (z - H_2)}}{\cos (\sigma_2^{(n)} H_2)} \left[\left\{ \frac{1}{M \frac{\partial}{\partial \lambda}(-\triangle)} \right\}_{\lambda=\lambda_n} \right]^{\frac{1}{2}} \tag{4.26}$$

and $\phi_1^{(n)} (z)$ is given by (4.11). From (4.21), (4.23) and (4.25), we obtain

$$I_{31} = \sum_{n=1}^{N} \phi_3^{(n)} (z)\, \phi_1^{n} (\zeta) - \int_{-\infty}^{\omega^2/\beta_3^2} \psi_3 (z, \lambda)\, \psi_1 (\zeta, \lambda)\, d\lambda, \tag{4.27}$$

where $\phi_1^{(n)}$, $\phi_3^{(n)}$, ψ_1, ψ_3 are given by (4.11), (4.26), (4.8) and (4.24) respectively.

All the other integrals can be manipulated in the same manner as in (*i*) – (*iii*). The final result is

$$I_{ij} = \lim_{R\to\infty} \frac{-1}{2\pi i} \oint_{|\lambda|=R} G_{ij}(z, \zeta ; \lambda)\, d\lambda$$

$$= \sum_{i=1}^{N} \phi_i^{(m)}(z)\, \phi_j^{(n)}(\zeta) - \int_{-\infty}^{\omega^2/\beta_3^2} \psi_i(z, \lambda)\, \psi_j(\zeta, \lambda)\, d\lambda, \tag{4.28}$$

(i, j = 1, 2, 3) with G_{ij} given by (3.5), (3.7), (3.8), (3.14), (3.16), (3.22).

From (4.1) and (4.28), we obtain the following representation of the delta function :

$$\delta(z-\zeta) = \sum_{n=1}^{N} \mu(\zeta)\, \phi^{(n)}(z)\, \phi^{(n)}(\zeta)$$

$$- \int_{-\infty}^{\omega^2/\beta_3^2} \mu(\zeta)\, \psi(z, \lambda)\, \psi(\zeta, \lambda)\, d\lambda, \tag{4.29}$$

where $\mu(\zeta)$ is given by (2.2),

$$\begin{aligned} \phi^{(n)}(z) &= \phi_1^{(n)}(z), \ -H_1 \leq z \leq 0, \\ &= \phi_2^{(n)}(z), \ 0 \leq z \leq H_2, \\ &= \phi_3^{(n)}(z), \ H_2 \leq z, \end{aligned} \tag{4.30}$$

are the normalized eigenfunctions, and

$$\begin{aligned} \psi(z, \lambda) &= \psi_1(z, \lambda), \ -H_1 \leqslant z \leq 0, \\ &= \psi_2(z, \lambda), \ 0 \leqslant z \leq H_2, \\ &= \psi_3(z, \lambda), \ H_2 \leq z \end{aligned} \tag{4.31}$$

are normalized improper eigenfunctions.

If $f(z)$ is of finite μ-norm over the interval $(-H_1, \infty)$, then the representation of $f(z)$ in terms of eigenfunctions $\{\phi^{(n)}(z)\}$ and improper eigenfunctions $\{\psi(z, \lambda)\}$ can be obtained on multiplying (4.29) by $f(\zeta)$ and integrating with respect to ζ from $-H_1$ to ∞. We get

$$\int_{-H_1}^{\infty} f(\zeta)\, \delta(z-\zeta)\, d\zeta = \sum_{n=1}^{N} \phi^{(n)}(z) \int_{-H_1}^{\infty} \mu(\zeta) f(\zeta)\, \phi^{(n)}(\zeta)\, d\zeta$$

$$- \int_{-\infty}^{\omega^2/\beta_3^2} \psi(\lambda, z)\, dz \int_{-H_1}^{\infty} \mu(\zeta)\, \psi(\zeta, \lambda)\, f(\zeta)\, d\zeta,$$

whence

$$f(z) = \sum_{n=1}^{N} f_n\, \phi^{(n)}(z) - \int_{-\infty}^{\omega^2/\beta_3^2} f_\lambda\, \psi(\lambda, z)\, dz, \qquad (4.32)$$

where

$$f_n = \langle f, \phi^{(n)} \rangle = \int_{-H_1}^{\infty} \mu(\zeta)\, f(\zeta)\, \phi^{(n)}(\zeta)\, d\zeta, \qquad (4.33)$$

and

$$f_\lambda = \langle f, \psi(\zeta, \lambda) \rangle = \int_{-H_1}^{\infty} \mu(\zeta)\, \psi(\zeta, \lambda)\, f(\zeta)\, d\zeta. \qquad (4.34)$$

In particular, if $f(z) = \phi^{(m)}(z)$ or $\psi(z, \lambda')$, then (4.32) – (4.35) yield the following orthonormality relations :

$$\int_{-H_1}^{\infty} \mu(z)\, \phi^{(m)}(z)\, \phi^{(n)}(z)\, dz = \delta_{mn} = \langle \phi^{(m)}, \phi^{(n)} \rangle,$$

$$1 \le m, n \le N \qquad (4.35a)$$

$$\int_{-H_1}^{\infty} \mu(z)\, \psi(z,\lambda)\, \psi(z,\lambda')\, dz = \delta(\lambda-\lambda') = \langle \psi(z,\lambda), \psi(z,\lambda') \rangle ,$$

$$-\infty < \lambda, \lambda' < \omega^2 / \beta_3^2 \qquad (4.35b)$$

$$\int_{-H_1}^{\infty} \mu(z)\, \phi^{(m)}(z)\, \psi(z,\lambda)\, dz = 0 = \langle \phi^{(m)}, \psi \rangle , \quad 1 \leq m \leq N,$$

$$-\infty < \lambda < \omega^2 / \beta_3^2 \qquad (4.35c)$$

5. Conclusion

We have obtained the spectral representation of the two-dimensional Love wave operator associated with monochromatic SH-waves in a structure comprising two homogeneous layers overlying a uniform half-space. The spectral representation enables us to tackle classes of problems associated with the transmission and reflection of Love waves at a horizontally discontinuous change either in elevation or in material properties of three-layered models, using the method based on an integral equation formulation together with the application of Schwinger-Levine variational principle as in Kazi [1978a, b] and Niazy and Kazi [1980].

Acknowledgements

The authors wish to thank Dr. David Willmer for several valuable discussions and suggestions, and also to the University of Petroleum and Minerals for providing excellent facilities.

REFERENCES

1. Ewing, W.M., Press, F. and Jardetsky, W.S., 1957. Elastic waves in layered media, McGraw-Hill.

2. Kazi, M.H., 1976. Spectral representation of the Love wave operator, Geophys. J.R. Astr. Soc., **47**, 225-249.

3. Kazi, M.H., 1978a. The Love wave scattering matrix for a continental margin (Theoretical), Geophys. J.R. Astr. Soc., **52**, 25-44.

4. Kazi, M.H., 1978b. The Love wave scattering matrix for a continental margin (Numerical), Geophys. J.R. Astr. Soc., **53**, 227-243.

5. Kazi, M.H., 1979. Transmission, reflection and diffraction of Love waves in an infinite strip with a surface step, J. Phys. A., Math. Gen., **12**, 1441-1455.

6. Kennett, B.L.N., 1981. On coupled seismic waves, Geophys. J.R. Astr. Soc., **64**, 91-114.

7. Niazy, A. and Kazi, M.H., 1980. On the Love wave scattering problem for welded layered quarter-spaces with applications, Bull. Seism. Soc. Am. **70**, 2071-2095.

The Punjab University
Journal of Mathematics
Vol. XIV—XV (1981-82), pp. 71-85

SPECIFICATION OF THE NON-INFORMATIVE PRIOR DISTRIBUTIONS IN THE BAYESIAN ANALYSIS OF THE ADAPTIVE EXPECTATIONS MODELS

By

NOOR MUHAMMAD LARIK

Statistics Department,
P.O. Box. 13217,
Al-Fateh University,
Tripoli, LIBYA.

Abstract.

The main problem in Bayesian analysis of the adaptive expectations model is the assignment of prior distributions of parameters when very little is known *a priori* about them. Various approaches have been suggested to assign prior distributions in such a situation. It is shown that these approaches are applicable to the parameters, β_i's and σ_u^2 in the adaptive expectations model but not to the parameter λ. Furthermore, it is shown that the assumption of taking implicitly uniform prior distribution for λ is not too bad for small n. However, for large n the prior distribution of λ is suggested a binomial prior with $P = \frac{\lambda+1}{2}$. $0 < P < 1$.

Introduction

An adaptive expectations model in general can be written as

$$y_t = \sum_{i=1}^{k} \beta_i \sum_{j=0}^{t-1} \lambda^i X_{i\,t-j} + m\lambda^t + u_t, \quad t = 1, 2, \ldots, n,$$

where, y_t is an endogenous variable, $X_{i\,t-j}$ is an exogenous variable ; m (the initial condition), β_i's and λ are unknown parameters such that, $-\infty < m, \beta_i < \infty$, $0 \leq \lambda < 1$ and u_t is the disturbance term and is necessarily assumed to be normally distributed with mean zero and common variance σ_u^2. This model can be represented in matrix notation as

$$Y = m\,\Lambda + CX\,\beta + U,$$

where,

$$\underset{l\times n}{Y^T} = (y_1, y_2, \ldots\ldots, y_n), \quad \underset{l\times n}{\Lambda^T} = (\lambda, \lambda^2, \ldots\ldots, \lambda^n),$$

$$\underset{l\times k}{\beta^T} = (\beta_1, \beta_2, \ldots\ldots, \beta_k), \quad \underset{l\times n}{U^T} = (U_1, U_2, \ldots\ldots, U_n)$$

and

$$\underset{n\times n}{C} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ \lambda & 1 & \ldots & 0 \\ \lambda^{n-1} & \lambda^{n-2} & \ldots & 1 \end{bmatrix}, \quad \underset{n\times k}{X} = \begin{bmatrix} X_{11} & X_{12} & \ldots\ldots & X_{1k} \\ X_{21} & X_{22} & \ldots\ldots & X_{2k} \\ X_{n1} & X_{n2} & \ldots\ldots & X_{nk} \end{bmatrix}$$

The maximum like lihood function of this model can be written [Rao (1965)] as

$$L(\theta \mid Y) = \frac{1}{(2\pi)^{n/2}\,(\sigma_u^2)^{n/2}} \cdot \exp\left\{ -\frac{1}{2\sigma_u^2} (Y - m\,\Lambda - CX^{\beta})^T\,(Y - m\,\Lambda - CX^{\beta}) \right\}$$

where

$$\theta^T = (m, \beta^T, \lambda^T, \sigma_u^2).$$

Now, in Bayesian analysis the estimation of parameters and inference is based on posterior probability density function of the parameters say $P_2(\theta \mid Y)$, which may be obtained by multiplying two input functions, the likelihood function $L(\theta \mid Y)$ and the prior density function, say, $P_1(\theta)$ as :

$$P_2(\theta \mid Y) \propto L(\theta \mid Y) \times P_1(\theta).$$

The main problem in the Bayesian analysis of the adaptive expectations model is to formulate the prior density $P_1(\theta)$ when no information is available about the concerned parameters. To carry on the Bayesian analysis in such a situation we examine the applicability of various approaches that have been suggested by Jefferys' (1961), Raiffa and Schlaifer (1961), Lindley (1965), Novick and Hall (1965), Jaynes (1968) and Lindley and Smith (1972).

We will see below that the above approaches for assigning prior distributions are applicable to β's and σ_u^2 in our case but not to the parameter λ.

JEFFERYS' APPROACH

First of all, we shall employ two well known rules of Jeffreys'. According to his first rule the prior distribution representing the state of ignorance for a parameter vector $\theta^T = (m, \beta^T, \lambda^T, \sigma_u^2)$ is taken to be

$$P_1(m, \beta, \lambda, \sigma_u^2) \propto \frac{1}{\sigma_u},$$

$$-\infty < m, \beta < \infty,$$

$$0 \leq \lambda < 1,$$

$$0 \leq \sigma_u < \infty,$$

where m, β, λ and $\log \sigma_u$ are assumed uniformly and independently distributed, (the implicit uniform prior for λ does not accord with Jeffreys' principle). Jeffreys' second rule of assigning priors representing a state of ignorance (i.e., priors obtained by taking the square root of the determinant of the information matrix) does not always give acceptable results. In our case the information matrix, say I (θ)

$$I(\theta) = \frac{1}{\sigma_u^2}$$

$$\begin{bmatrix} \Lambda^T \Lambda & \Lambda^T CX & . \; m\Lambda^T \frac{\partial \Lambda}{\partial \lambda} + \Lambda^T DX\beta & . \; 0 \\ & X^T C^T CX. & m X^T C^T \frac{\partial \Lambda}{\partial \lambda} + X^T C^T DX\beta & . \; 0 \\ \cdots & \cdots & \cdots & \cdots \\ & & \vdots \; (m\frac{\partial \Lambda}{\partial \lambda} + DX\beta)^T (m\frac{\partial \Lambda}{\partial \lambda} + DX\beta) & \vdots \; 0 \\ & & & 2n \end{bmatrix}$$

(Where, I (θ) is obtained by taking minus the expected values of the second order partial derivative of the log likelihood function of the model with respect to parameters and then :

Substituting $\frac{m}{\sqrt{n}} = \alpha_o$, $D = \frac{\partial C}{\partial \lambda}$, and deleting the n's), depends on the data and involves a large number of parameters hence, it would be impracticable to use Jeffreys' second rule. Taking invariance for each individual parameter is also impracticable since λ is not separable from the X's. However, if we consider a more general model, of which an adaptive expectation model is a special case, of the form

$$y_t = \theta_1 y_{t-1} + \sum_{i=1}^{k} \beta_i x_{it} + u_t - \lambda u_{t-1}$$

and assume y_o fixed, u_o not fixed and apply an orthogonal transformation used by Pesaran [1973] to the variance-covariance matrix σ_u^2 G say, where

$$G = \begin{bmatrix} 1+\lambda^2 & -\lambda & 0 \ldots\ldots 0 \\ -\lambda & 1+\lambda^2 & -\lambda \ldots\ldots 0 \\ 0 & & 0 \\ \cdot & & \\ \cdot & & \\ \cdot & & -\lambda \\ 0 & -\lambda & 1+\lambda^2 \end{bmatrix}$$

If θ^T now $= (\theta_1, \beta, \lambda, \sigma_u^2)$ we obtain a simple form of the information matrix, that is,

$$I_1(\theta) = \begin{bmatrix} \dfrac{X^T \triangle^{-1} X}{\sigma_u^2} & 0 & 0 \\ 0 & 2\sum_{t=1}^{n} \dfrac{K_{2t^2}}{K_{1t^2}} & \dfrac{2}{\sigma_u^2}\sum_{t=1}^{n} \dfrac{K_{2t}}{K_{1t}} \\ 0 & \dfrac{2}{\sigma_u}\sum_{t=1}^{n} \dfrac{K_{2t}}{K_{1t}} & \dfrac{2n}{\sigma_u^2} \end{bmatrix}$$

where,

$$K_{1t} = \lambda^2 - 2\lambda \cos\left(\frac{\pi t}{n+1}\right) + 1,$$

$$K_{2t} = \lambda - \cos\left(\frac{\pi t}{n+1}\right),$$

$$\triangle = TGT^T, \quad \triangle^{-1} = TG^{-1}T^T$$

$$G = T^T \triangle T, \quad T^T = \left(\frac{2}{n+1}\right)^{\frac{1}{2}} (t_1, t_2, \ldots\ldots, t_n),$$

$$t_j^T = \left(\sin \frac{j\pi}{n+1}, \sin \frac{2j\pi}{n+1}, \ldots\ldots\ldots, \sin \frac{nj\pi}{n+1} \right),$$

$$\Delta_{ij} = \lambda^2 - 2\lambda \cos\left(\frac{j\pi}{n+1}\right) + 1 \qquad i = j$$

$$= 0 \qquad i \neq j$$

$$i, j = 1, \ldots\ldots, n,$$

$$|G| = \frac{1 - \lambda^{2n+2}}{1 - \lambda^2} = \prod_{t=1}^{n} \left(\lambda^2 - 2\lambda \cos\left(\frac{\pi t}{n+1}\right) + 1\right)$$

If we now apply Jeffreys' second rule we can obtain priors which seem to be practicable. Assuming β to be independently distributed of λ and σ_u we have

$$P_1(\beta, \lambda, \sigma_u) \propto P_1(\beta)\ P_1(\lambda, \sigma_u).$$

$P_1(\beta)$ is assumed uniformly and independently distributed as

$$P_1(\beta) \propto 1$$

Where, $I_1(\theta)$ is obtained similarly as $I(\theta)$. The log likelihood function in this case is

$$\text{Log } L = -\frac{n}{2}\log 2\pi - n \log \sigma_u - \tfrac{1}{2}\log |G| - \frac{1}{\sigma_u^2}(Y - x^\beta)\ T\ .\ (Y - x^\beta).$$

$$P_1(\lambda, \sigma_u) \propto \begin{vmatrix} 2\sum_{t=1}^{n} \frac{K_{2t}^2}{K_{1t}^2} & \frac{2}{\sigma_u}\sum_{t=1}^{n} \frac{K_{2t}}{K_{1t}} \\ \frac{2}{\sigma_u}\sum_{t=1}^{n} \frac{K_{2t}}{K_{1t}} & \frac{2n}{\sigma_u^2} \end{vmatrix}^{\frac{1}{2}}$$

We now have a choice. We can treat each parameter separately. This is equivalent to making our posterior inferences invariant under reparameterizations of the form

$$\overset{*}{\lambda} = \overset{*}{\lambda}(\lambda),$$

$$\overset{*}{\sigma} = \overset{*}{\sigma}(\sigma),$$

but not of the form $(\overset{*}{\lambda}, \overset{*}{\sigma}) = \phi(\lambda, \sigma)$, i.e. under joint transformation. Separate treatment yields

$$P_1(\lambda) \propto \left| \sum_{t=1}^{n} \frac{K_{2t}^2}{K_{1t}^2} \right|^{\frac{1}{2}}$$

$$P_1(\sigma_u) \propto \left| \frac{n}{\sigma_u^2} \right|^{\frac{1}{2}}$$

Alternatively, we can treat λ, σ_u jointly to have

$$P_1(\lambda, \sigma_u) \propto \frac{n^{\frac{1}{2}}}{\sigma_u} \left\{ \sum_{t=1}^{n} \frac{K_{2t}^2}{K_{1t}} - n \left[\overline{\frac{K_{2t}}{K_{1t}}} \right]^2 \right\}^{\frac{1}{2}}$$

and thus

$$P_1'(\lambda) \propto \left\{ \sum_{t=1}^{n} \left[\frac{K_{2t}}{K_{1t}} \right]^2 - n \left[\overline{\frac{K_{2t}}{K_{1t}}} \right]^2 \right\}^{\frac{1}{2}},$$

where

$\left\{ \overline{\frac{K_{2t}}{K_{1t}}} \right\}$ is mean of the ratio $\left\{ \frac{K_{2t}}{K_{1t}} \right\}$.

We have computed $P_1(\lambda)$ and $P_1'(\lambda)$ for $n = 2, 10, 20, 30,$ and $40\ (-1 \leq \lambda \leq 1)$. These are plotted in Figs. (1) and (2).

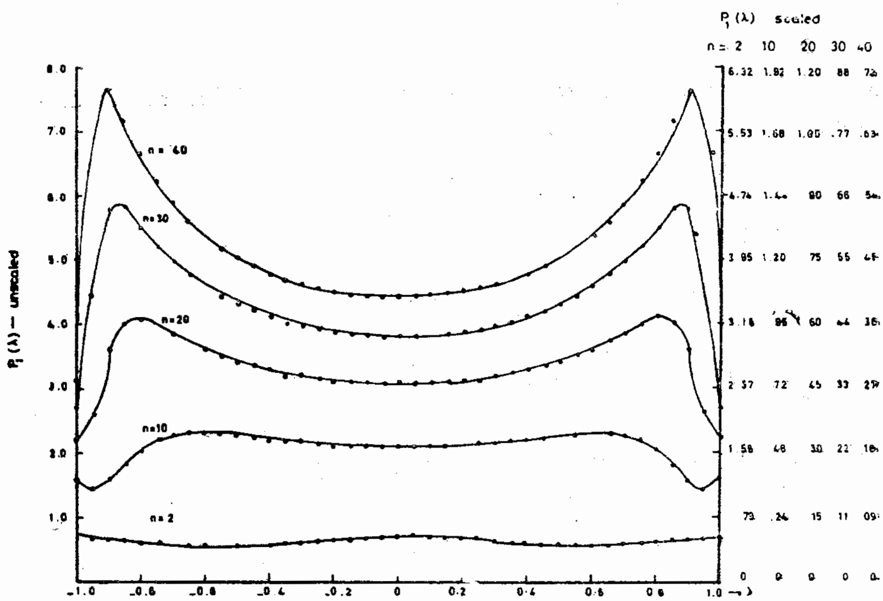


Fig: 1 Non-informative prior distribution $P_1(\lambda)$

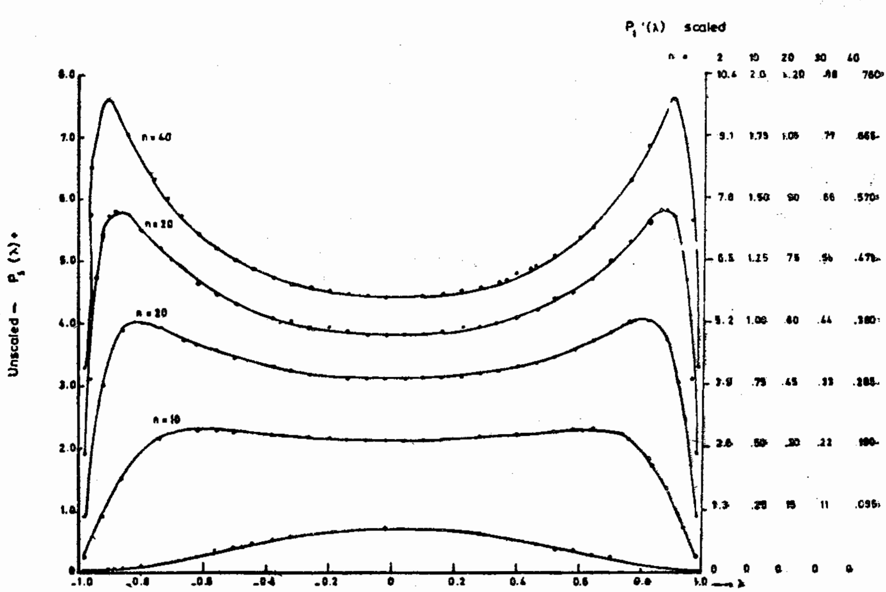


Fig: 2 Non-informative prior distribution $P_1'(\lambda)$

This set of assumptions : y_0 fixed, u_0 not fixed (as var $(u_1 - \lambda u_0)$ $= ((1+\lambda^2)\sigma_u^2)$ is mildly contradictory to the A.E. model as it requires $y_0 \neq m + u_0$, where m is non-stochastic. However, the A.E. model is a special case.

Now we assume u_0 also fixed, and zero. This is to obtain $E(w_t) = 0$ where, $w_t = u_t - \lambda u_{t-1}$. Under these constraints the variance-covariance matrix σ_u^2 G given above becomes σ_u^2 V, where,

$$V = \begin{bmatrix} 1 & -\lambda & 0 \ldots\ldots 0 & \\ -\lambda & 1+\lambda^2 & -\lambda & \\ 0 & & & \\ \vdots & & & -\lambda \\ \vdots & & -\lambda & 1+\lambda^2 \end{bmatrix}$$

$$= C^{-1}(C^{-1})^T$$

$$C^{-1} = \begin{bmatrix} 1 & 0 \ldots\ldots 0 & \\ -\lambda & 1 \ldots\ldots 0 & \\ 0 & & \\ \vdots & & \\ \vdots & -\lambda & 1 \end{bmatrix}$$

and $V^{-1} = C^T C$,

$$C^T = \begin{bmatrix} 1 & \lambda & \ldots\ldots & \lambda^{n-1} \\ 0 & & & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ 0 & & & 1 \end{bmatrix}$$

the i-jth element of V^{-1} say v^{ij} is

$$v^{ij} = \lambda^{j-i} \frac{(1-\lambda^{2(n+1-j)})}{(1-\lambda^2)} \qquad j \geq i$$

$$= \lambda^{i-j} \frac{1-\lambda^{2(n+1-i)})}{(1-\lambda^2)} \qquad i \geq j$$

$$i, j = 1, 2, \ldots\ldots, n.$$

The information matrix in this case $I_2(\theta)$, say, is

$$I_2(\theta) = \begin{bmatrix} \frac{X^{*T} V^{-1} X^*}{\sigma_u^2} & 0 & 0 \\ 0 & \frac{2n}{1-\lambda^2} - \frac{2(1-\lambda^{2n})}{(1-\lambda^2)^2} & 0 \\ 0 & 0 & \frac{2n}{\sigma_u^2} \end{bmatrix},$$

The information matrix $I_2(\theta)$ looks much simpler than $I_1(\theta)$. Applying Jeffreys' second rule we obtain the non-informative prior for λ say $P_1''(\lambda)$ as

$$P_1''(\lambda) \propto \left| \frac{n}{1-\lambda^2} - \frac{(1-\lambda^{2n})}{(1-\lambda^2)^2} \right|^{\frac{1}{2}}$$

$P_1''(\lambda)$ has been computed for $n = 2, 10, 20, 30, 40$ and plotted in Fig. 3. From the figures it looks that the assumption of taking implicity uniform prior is not too bad, since prior distributions are fairly flat within the range $-.8 \leq \lambda \leq .8$, particularly, see for example, $P_1''(\lambda)$ plotted in Fig. 3.

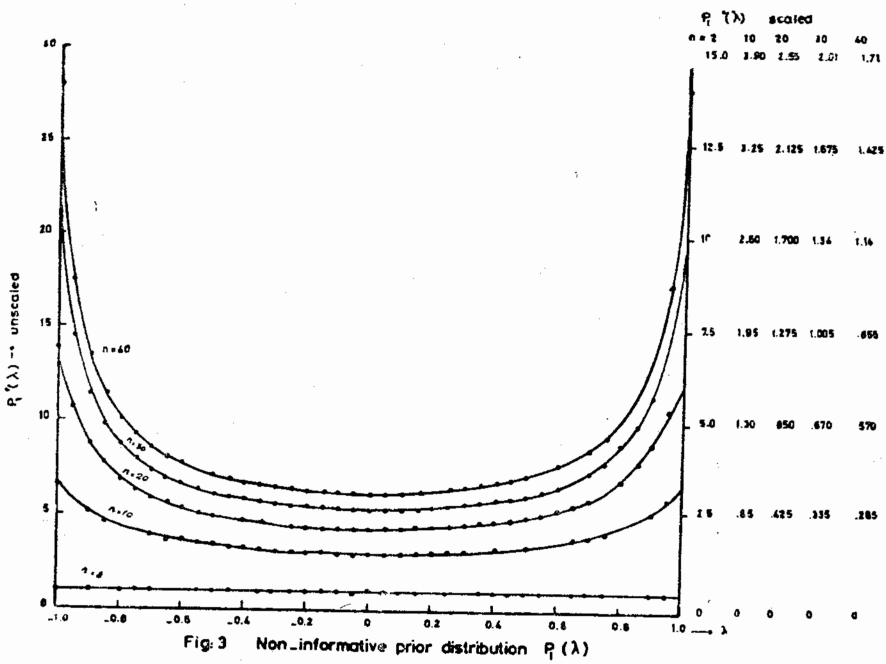


Fig. 3

From these figures it appears that the prior distribution for λ depends on n, for example, for large n.

$$P_1''(\lambda) \propto \left| \frac{n}{1-\lambda^2} - \frac{1-\lambda^{2n}}{(1-\lambda^2)^2} \right|^{\frac{1}{2}} \text{ behaves like } \frac{1}{\sqrt{1-\lambda^2}}$$

In econometrics we are likely to have n large, so that the prior suggested is $\frac{1}{\sqrt{1-\lambda^2}}$. This is exactly the binomial prior with $P = \frac{\lambda+1}{2}, 0 < P < 1$.

OTHER APPROACHES

The "indifference rule" of Novick and Hall (1965) for assigning prior distributions in a state of ignorance requires the existence of the family of natural conjugate distributions ; such a family of distributions may be obtained if a sufficient statistic exists. In our case, we do not have a sufficient statistic hence we can not construct a family of conjugate distributions. This suggests that indifference rule is not applicable.

Strictly speaking, one can always write

$$L(\underline{x} \setminus \theta) = L(\underline{x} \setminus \theta) \times l$$

and the set of observations $\underline{x}$ becomes a set of sufficient statistics.

However, what we require is a set of sufficient statistics of fixed dimension, Raiffa and Schlaifer (1961, p. 44) and this is not available. A set of sufficient statistics of fixed dimension would be available if we assume λ to be a known parameter, in which case a family of conjugate priors for β, σ_u^2 can be constructed and the indifference rule can be applied. The suggestion of letting $\sigma_u \rightarrow \infty$ in conjugate distribution made by Raiffa and Schlaifer (1961) and Lindley (1965) for assigning prior distribution in a state of ignorance will also be workable if λ is known.

Jaynes (1968) employs the principle of maximum entropy to allocate priors for discrete parameter spaces. For continuous parameter spaces this is not directly applicable. Accordingly, he also employs considerations group invariance. Where the number of parameters in the transformations is equal to the number of parameters in the prior, a unique prior is specified : there is no need to invoke maximum entropy. He employs the device of invariance under a group of transformations to derive the usual non-informative priors for the Normal and Poisson cases, and $P^{-1}(1-P)^{-1}$ for the Binomial, but his choice of transformation appears adhoc. For

Jaynes' proposal, there does not seem to be any immediately appealing choice of transformations groups for λ in our case.

CONCLUSIONS

The application of Jeffreys' rules for assigning prior distributions to parameters of the adaptive expectations models requires generalisation of the model. The other approaches suggested for assigning prior distributions representing prior ignorance are applicable to β's and σ_u^2 in the case of adaptive expectations model but not to λ. However in the case of a smaller number of observations a uniform prior distribution for λ can be used, whereas for large n the prior distribution of λ is suggested a binomial with parameter $P = \frac{\lambda + 1}{2}$, $0 < P < 1$.

REFERENCES

1. Box, G.E.P. and Tiac, G.C. (1973), Bayesian Inference in Statistical Analysis, Addison Wesley.
2. Dawid, A.P., Stone, M. and Zidek, J.V. (1973), Marginalization Paradoxes in Bayesian and Structural Inference, (with Discussion) Jr. Royal Statistical Society (Ser. B), 35, 189-233.
3. Feller, W. (1971), An Introduction to Probability Theory and Its Applications, II, 2nd Edn. John Wiley and Sons, Inc.
4. Hewitt, E. and Savage, L.J. (1955), Symmetric Measures on Cartesian products, Trans, Amer. Math. Soc. 80, 470-501.
5. Hartigan, J. (1964), Invariant Prior Distributions, Ann. Math. Statistics., 35, 836-845.

6. Jaynes, E.T. (1968), Prior Probabilities, IEEE Trans. System Science and Cybernetics SSC-4, 227.

7. Jeffreys, H. (1961), Theory of Probability, Third Edition, Clarendon Press, Oxford.

8. Kendall, M.C. and Stuart, A. (1961), The Advanced Theory of Statistics 2, Charles Griffin and Company Limited, London.

9. Lindley, D.V. (1965), Probability and Statistics : Inference, 2, Cambridge at the University Press.

10. Lindley, D.V. and Smith, A.F.M. (1972), Bayes Estimates for the Linear Model, (with Discussion). Journal of the Royal Statistical Society, (Ser. B), 34, 1-41.

11. Lindley, D.V. (1971, 1972), Bayesian Statistics, A Review SIAM Philadphia, PA.

12. Novick, M.R. and Hall, W.H. (1965), A Bayesian Indefference Procedures, Jr. of Royal Statistical Association, 60, 1104-1117.

13. Novick, M.R. (1969), Multiparameter Bayesian Indefference Procedures, Jr. of Royal Statistical Society (series **B**), **31**, 29-64.

14. Pesaran, M.H. (1973), Exact Maximum Likelihood Estimation of a Regression Equation with a First-Order Moving Average Error, The Review of Economic Studies, 529-533.

15. Plackett, R.L. (1966), Current Trends in Statistical Inference. J. Royal Stats. Society (B), 249-267.

16. Raiffa, H. and Schlaifer, R. (1961), Applied Statistical Decision Theory, Boston, Harvard Business School.

17. Thornber, H. (1967), Finite Sample Monte Carlo Studies : An Autoregressive Illustration, Journal, American Statistical Association, **62**, 801-818.

18. Zellner, A. (1971), An Introduction to Bayesian Inference in Econometrics, John Wiley and Sons, Inc, New York, London Sydney Toronto.

19. Zellner, A. and Williams, A.D. (1973), Bayesian Analysis of the Federal Reserve – MIT – Penn Model's Almon Lag Consumption Function, Journal of Econometrics, 1, 267-299.

20. Rae, (1965), Linear Statistical Inference and its Applications, John Wiley and Sons, Inc.

The Punjab University
Journal of Mathematics
Vol. XIV—XV (1981-82), pp. 86-108

ON THE DISTRIBUTION OF PARTIAL AND MULTIPLE CORRELATION COEFFICIENTS WHEN SAMPLING FROM A MIXTURE OF TWO MULTIVARIATE NORMALS

By

HAYAT M. AWAN

Statistics Department,
Bahauddin Zakariya University, Multan, Pakistan
and M.S. Srivastava
University of Toronto, (Canada)

Abstract.

In this paper the null and non-null distributions of partial and multiple correlation coefficients are derived when the sample is taken from a mixture of two p-component multivariate normal distributions with mean vectors $\underline{\mu}_1$ and $\underline{\mu}_2$ respectively and common covariance matrix Σ. Some special cases are also given.

Keywords and phrases :

Non-central wishart, non-centrality matrix, standardized distance, doubly non-central F distribution, Invariance.

AMS Subject Classification : Primary 62H 10, Secondary 62H 20.

1. Introduction

The study of statistical relation between variables occupies an important place in the subject matter of statistics. One aspect of this study is the theory of correlation based on the various correlation coefficients which was mathematically developed by Fisher [1915].

The very extensive study of the simple correlation coefficient contrasts sharply with the little attention given to the partial and multiple correlation coefficients. The distributional problem of partial and multiple correlation coefficients has been left virtually as it was found by Fisher in the normal case. This appears to be due to mathematical difficulty involved in non-normal cases.

In this paper we derive the null and non-null distributions of the partial correlation $r_{12\,.\,34\,\ldots\,p}$ and the multiple correlation R^2 when a sample of size N is drawn from a population with probability density function (p.d.s), given by

$$f(\mathbf{x}) = \lambda\,\phi\,(\mathbf{x}\,;\underline{\mu}_1,\Sigma)+(1-\lambda)\,\phi\,(\mathbf{x}\,;\underline{\mu}_2\,,\Sigma),\ 0 \le \lambda \le 1 \quad (1.1)$$

where $\phi\,(\mathbf{x}\,;\underline{\mu},\Sigma)$ denotes the pds of multivariate normal distribution with mean vector $\underline{\mu}$ and covariance matrix Σ and λ stands for the mixing proportion (contamination, $1-\lambda$).

It should be mentioned that G.A. Baker (1932) was the first to derive the distribution of Student ; t-statistic for a sample of two items from a composition of two normal functions with different means. The distribution of the simple correlation coefficient (null and non-null) has been derived by Srivastava and Awan [1980]. The posterior distribution of the location parameter and the effect of shift on location from such a model has been carried out by Awan and Srivastava [1980]. The null distribution of Hotelling's T^2 has been used by Srivastava and Awan [1982] to study the robustness of Hotelling's T^2 test. A numerical study of these results is planned for a subsequent publication.

2. Partial Correlation

Let $\mathbf{x}_1\,,\mathbf{x}_2\,,\ldots,\mathbf{x}_N$ be a random sample of size N on a p-compenent random vector $\mathbf{x}=(\mathbf{x}_1\,,\mathbf{x}_2\,,\ldots,x_p)$ with pds given by (1.1). Let the sample mean and the sample covariance denoted by $\mathbf{x}$ and $\mathbf{S}$ respectively be given by

$$\bar{\mathbf{x}} = N^{-1} \sum_{i=1}^{N} \mathbf{x}_i \text{ and } S = \sum_{i=1}^{N} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})', \qquad (2.1)$$

We partition S as

$$S = \begin{bmatrix} s_{11} & s_{12} & \mathbf{S}_{13}' \\ s_{12} & s_{22} & \mathbf{S}_{23}' \\ \mathbf{S}_{13} & \mathbf{S}_{23} & \mathbf{S}_{33} \end{bmatrix} \qquad (2.2)$$

where $s_{11} : 1 \times 1$, $s_{12} : 1 \times 1$, $s_{22} : 1 \times 1$, $\mathbf{S}_{13} : (p-2) \times 1$, $\mathbf{S}_{23} : (p-2) \times 1$ and $\mathbf{S}_{33} : (p-2) \times (p-2)$, the sample partial correlation coefficient $r_{12.34\ldots p}$ between $\mathbf{x}_1$ and $\mathbf{x}_2$ given $\mathbf{x}_3 = (\mathbf{x}_3, \ldots, \mathbf{x}_p)$ is defined as

$$r_{12.34\ldots p} = \frac{\text{cov}(\mathbf{x}_1, \mathbf{x}_2 \mid \mathbf{x}_3)}{\{\text{var}(\mathbf{x}_1 \mid \mathbf{x}_3)\, \text{var}(\mathbf{x}_2 \mid \mathbf{x}_3)\}^{\frac{1}{2}}} \qquad (2.3)$$

$$= \frac{s_{12} - \mathbf{S}_{13}' \mathbf{S}_{33}^{-1} \mathbf{S}_{23}}{\{(s_{11} - \mathbf{S}_{13}' \mathbf{S}_{33}^{-1} \mathbf{S}_{13})(s_{22} - \mathbf{S}_{23}' \mathbf{S}_{33}^{-1} \mathbf{S}_{23})\}^{\frac{1}{2}}}$$

The population partial correlation co-efficient between $\mathbf{x}_1$ and $\mathbf{x}_2$ gjven $\mathbf{x}_3$ is defined as

$$\rho_{12.34\ldots p} = \frac{\sigma_{12} - \underline{\sigma}_{13}' \Sigma_{33}^{-1} \underline{\sigma}_{23}}{\{(\sigma_{11} - \underline{\sigma}_{13}' \Sigma_{33}^{-1} \underline{\sigma}_{13})(\sigma_{22} - \underline{\sigma}_{23}' \Sigma_{33}^{-1} \underline{\sigma}_{23})\}^{\frac{1}{2}}} \qquad (2.4)$$

where Σ is partitioned as in S, that is,

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \underset{-}{\sigma}'_{13} \\ \sigma_{12} & \sigma_{22} & \underset{-}{\sigma}'_{23} \\ \underset{-}{\sigma}_{13} & \underset{-}{\sigma}_{23} & \Sigma_{33} \end{bmatrix} \tag{2.5}$$

If $\mathbf{x} \sim N_p(\underset{-}{\mu}, \Sigma)$, then it is well known, see for example Srivastava and Khatri [1979] or Anderson [1958] that the distribution of $r_{12\,.\,34\,\ldots\,p}$ is just that of ordinary correlation (in the bivariate case) on $(N - p + 1)$ degrees of freedom (hereafter called d.f.).

In this section we derive the null and non-null distribution of $r_{12\,.\,34\,\ldots\,p}$ when a sample of size N is drawn from a population with pdf given in (1.1).

The partial correlation co-efficient is invariant under the group G of transformation g, given by

$$gX = (x_1 + a_1, x_2 + a_2, \mathbf{x}_3 + \mathbf{a}_3)\,\lambda \begin{bmatrix} g_{11} & 0 & \underset{-}{0}' \\ 0 & g_{22} & \underset{-}{0}' \\ \underset{-}{0} & \underset{-}{0} & G_{33} \end{bmatrix}$$

where $g_{11} \neq 0$, $g_{22} \neq 0$, $G_{33} : (p-2) \times (p-2)$ is a nonsingular matrix and $a_1 : 1 \times 1$, $a_2 : 1 \times 1$, $\mathbf{a}_3 : (p-2) \times 1$ are real vectors.

We assume without any loss of generality that

$\underset{-}{\mu}_1 = 0$, $\underset{-}{\mu}_2 = \underset{-}{\delta}$ and

$$\Sigma = \begin{bmatrix} 1 & \beta_1 & \underset{-}{\beta}'_2 \\ \beta_1 & 1 & \underset{-}{\beta}'_3 \\ \underset{-}{\beta}_2 & \underset{-}{\beta}_3 & 1_{(p-2)} \end{bmatrix} \tag{2.6}$$

For fixed k, the sample covariance matrix can be written as

$$S = \sum_{i=1}^{N} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})',$$

$$= \sum_{i=1}^{k} (\mathbf{x}_i - \bar{\mathbf{x}}_1)(\mathbf{x}_i - \bar{\mathbf{x}}_1)',$$

$$+ \sum_{i=k+1}^{N} (\mathbf{x}_i - \bar{\mathbf{x}}_2)(\mathbf{x}_i - \bar{\mathbf{x}}_2)', \quad (2.7)$$

$$+ \frac{k(N-k)}{N} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)',$$

where, $\bar{\mathbf{x}}_1 = k^{-1} \sum_{i=1}^{k} \mathbf{x}_i$ and $\bar{\mathbf{x}}_2 = (N-k)^{-1} \sum_{i=k+1}^{N} \mathbf{x}_i$

Then for fixed k, S has a noncentral wishart distribution $W_p(\Sigma, n, \Omega_k)$ with $n = N - 1$ d. f. and non-centrality matrix

$$\Omega_k = N^{-1} k(N-k) \; \Sigma^{-1} \underline{\delta}\, \underline{\delta}'$$

which is of rank one. Hence the pdf of the sample covariance matrix S is given by

$$f_N(S) = \sum_{k=0}^{N} \binom{N}{k} \lambda^k (1-\lambda)^{N-k} f_N^k(S) \quad (2.7\text{ A})$$

where the sample is taken from (1.1) with $\underline{\mu}_1 = 0$, $\underline{\mu}_2 = \underline{\delta}$ and Σ is given in (2.6), and $f_N^k(S)$ denote the non-central wishart pdf $W_p(\Sigma, n, \Omega_k)$.

From (2.7) we can write

$$S = YY' = \begin{bmatrix} Y_1' \\ \cdots \\ Y_2' \end{bmatrix} (Y_1 : Y_2) \quad (\;.8)$$

where Y is an $p + n$ matrix with $n = (N - 1)$, $Y_1 : n \times 2$ and $Y_2 : n \times (p - 2)$. The n columns of Y are independently distributed with common covariance matrix Σ. The mean of the first column random vector is $c_k \underline{\delta}$ where $c_k = (k (N - k) / N)^{\frac{1}{2}}$ and the means of the remaining $(n - 1)$ column random vectors are zero.
In terms of Y's

$$S_{1.2} = Y_1' \left(1 - Y_2 (Y_2' Y_2)^{-1} Y_2' \right) Y_1 \tag{2.9}$$

where

$$S_{1.2} = \begin{bmatrix} s_{11} - S_{13}' S_{33}^{-1} S_{13} & s_{12} - S_{13}' S_{33}^{-1} S_{23} \\ s_{12} - S_{13}' S_{33}^{-1} S_{23} & s_{22} - S_{23}' S_{33}^{-1} S_{23} \end{bmatrix} \tag{2.10}$$

and, correspondingly, the population matrix $\Sigma_{1.2}$ is given by

$$\Sigma_{1.2} = \begin{bmatrix} 1 - \underline{\beta}_2' \underline{\beta}_2 & \beta_1 - \underline{\beta}_2' \underline{\beta}_3 \\ \beta_1 - \underline{\beta}_2' \underline{\beta}_3 & 1 - \underline{\beta}_3' \underline{\beta}_3 \end{bmatrix} \tag{2.11}$$

Then by definition, the partial correlation coefficients $r_{12\,.\,34\,\ldots\,p}$ and $\rho_{12\,.\,34\,\ldots\,p}$ are just equal to the simple correlation coefficient obtained from $S_{1.2}$ and $\Sigma_{1.2}$ respectively.

Let $Y_1' = (\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_n)$, $Y_2' = (\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n)$ and $\underline{\delta}' = (\underline{\theta}_1', \underline{\theta}_2')$ where $\underline{\theta}_1' = (\delta_1, \delta_2)$ and $\underline{\theta}_2' = (\delta_3, \delta_4, \ldots, \delta_p)$

Note that

$$\eta_1' = E(Y_1') = c_k(\theta_1\ \underline{0}, \ldots, \underline{0}) \text{ and}$$

$$\eta_2' = E(Y_1') = c_k(\underline{\theta}_2\ \underline{0}, \ldots, \underline{0}) \tag{2.12}$$

First, we derive the distribution of $S_{1.2}$ for fixed k. The conditional distribution of Y_1' given R_2' and fixed k is given by

$$N_{2,n}\left(\eta_1' + \begin{bmatrix} \underline{\beta}_2' \\ \\ \underline{\beta}_3' \end{bmatrix} (Y_2' - \eta_2'),\ \Sigma_{1.2},\ 1\right) \tag{2.13}$$

Where $N_{2,n}(\underline{\mu}, \Sigma)$ is a multivariate normal with mean $\underline{\mu}$ and covariance matrix Σ.

Hence, given Y_2, $S_{1.2}$ has noncentral wishart distribution with $(n - p + 2)$ d. f. and noncentrality matrix

$$\Omega_k^* = \Sigma_{1.2}^{-1}\left[\eta_1' + \begin{bmatrix} \underline{\beta}_2' \\ \\ \underline{\beta}_3' \end{bmatrix} (Y_2' - \eta_2')\right]$$

$$(I - Y_2\,(Y_2')^{-1}\,Y_2')\left\{\eta_1 + (Y_2 - n_2)\,(\underline{\beta}_2, \underline{\beta}_3)\right\}$$

Let $$\tilde{\underline{\delta}}_k = c_k\left(\underline{\theta}_1 - \begin{bmatrix} \underline{\beta}_2' \\ \\ \underline{\beta}_3' \end{bmatrix} \underline{\theta}_2\right) \text{ and } B = \sum_{i=2}^{n} \underline{v}_i\,\underline{v}_i' \tag{2.14}$$

then $B \sim W_{p-2}(I, n-1)$ and $Y'_2 Y_2 = \underset{\sim}{v}_1 \underset{\sim}{v}'_1 + B$.

Hence

$$\Omega^*_k = \tilde{\delta}_k (I - \underset{\sim}{v}'_1 (\underset{\sim}{v}'_1 + B)^{-1} \underset{\sim}{v}_1) \tilde{\delta}_k \Sigma^{-1}_{1.2}$$

$$= \tilde{\underset{\sim}{\delta}}_k \tilde{\underset{\sim}{\delta}}'_k / (1 + \underset{\sim}{v}'_1 B^{-1} \underset{\sim}{v}_1) . \Sigma^{-1}_{1.2} \tag{2.15}$$

Since $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_n$ are independently distributed, with

$\underset{\sim}{v}_1 \sim N_{(p-2)}(c_k \underset{\sim}{\theta}_2, I)$ and $\underset{\sim}{v}_j \sim N_{(p-2)}(\underset{\sim}{0}, I), j=2, 3,\ldots, n,$

we find

$$\frac{n-p+2}{p-2} \underset{\sim}{v}'_1 B^{-1} \underset{\sim}{v}_1 \sim F_{p-2, n-p+2}(\gamma^2_k), \tag{2.16}$$

where $F_{p-2, n-p+2}(\gamma^2_k)$ denotes a non-central F distribution with $(p-2, n-p+2)$ d. f. and non-centrality parameter γ^2_k which is given by

$$\gamma^2_k = c^2_k \underset{\sim}{\theta}'_2 \underset{\sim}{\theta}_2$$

Hence we get the following,

Theorem 1.

Let $\mathbf{x}_1, \mathbf{x}, \mathbf{x}_2, \ldots, \mathbf{x}_N$ be independently distributed with common pdf $f(\mathbf{x})$ given by

$$f(\mathbf{x}) = \lambda \phi(\mathbf{x}, 0, \Sigma) + (1-\lambda) \phi(\mathbf{x}, \delta, \Sigma), \; 0 \leq \lambda \leq 1$$

where $\phi(\mathbf{x}, \underset{\sim}{\mu}, \Sigma)$ denotes a p-variate normal pdf with mean vector

$\underline{\mu}$ and covariance matrix Σ. Then the conditional distribution of $S_{1.2}$ in (2.10) for given k and Y_2 or $F_{p-2,\, n-p+2}\, (\gamma_k^2$ in (2.16), is a non-central wishart, distribution $w_2\, (\Sigma_{1.2}\, ,\, n-p+2,\, \Omega_k^*$, where

$$\Omega_k^* = \Sigma_{1.2}^{-1}\, \underset{\sim}{\tilde{\delta}}_k\, \underset{\sim}{\tilde{\delta}}_k \,/\, \{1 + \frac{p-2}{n-p+2}\, F_{p-2,\, n-p+2}\, (\gamma_k^2\,))$$

and $\Sigma_{1.2}$ is given (2.11).

Since the partial correlation coefficient is just equal to the simple correlation coefficient obtained from $S_{1.2}$ the distribution of $x = r_{12.34\, ,\, \ldots\, ,\, p}$ for fixed k can be written as

$$f_k\, (x \,/\, \rho^*) = \int_0^{\infty} \{\exp\, (-\, (a_1^2\, (k) - 2\, \rho^* a_1\, (k)\, a_2\, (k) + a_1^2\, (k)\,) \,/\, 2\, (1-\rho^{*2})\,)$$

$$\cdot\, \frac{(1-\rho^{*2})^{\frac{N-p+1}{2}}}{\sqrt{\pi}\, \Gamma\, ((N-p)\, /\, 2)} \sum_{m=0}^{\infty} \frac{(\, (a_2\, (k) - \rho^* a_1\, (k)\, /\, 2)\,)^{2m}}{m\, !\, \Gamma\, (\frac{N-p-1}{2} + m + 1)} \sum_{t=0} \frac{(2\, \rho^*\,)^t}{t\, !}$$

$$\cdot \sum_{i=0}^{m} \binom{m}{i} \left[\frac{2\, (a_1\, (k) - \rho^* a_2\, (k)\,)}{a_2\, (k) - \rho^* a_1\, (k)} \right]^i x^{i+t} \left(1 - x^2\right)^{\frac{N-p-2}{2}}$$

$$\cdot \sum_{j=0}^{m-i} \binom{m-i}{j} \left[\frac{a_1\, (k) - \rho^* a_2\, (k)}{a_2\, (k) - \rho^* a_1\, (k)} \right]^{2j}$$

$$.\ \Gamma\, (\, (N-p+1+i+t+2j\,)\, /\, 2)$$

$$.\ \Gamma \left(\frac{N-p+1+2m+t-i-2j}{2} \right).\ h\, (F)\, \}\, dF. \qquad (2.17)$$

where

$$a_1^2\, (k) = c_k^2\, (\delta_1 - \underset{-2}{\beta'}\, \underset{-3}{\delta}\,)^2 \,/\, (1 + \frac{p-2}{n-p+2}\, F\, (\gamma_k^2\,)\,),$$

$$a_2^2(k) = c_k^2 (\delta_2 - \underline{\beta}_3' \underline{\delta}_3)^2 / (1 + \frac{p-2}{n-p+2} F(\gamma_k^2)),$$

$$\underline{\delta}' = (\delta_1, \delta_2, \underline{\delta}_3) ; \underline{\delta}_3 = (\delta_3, \delta_4, \ldots, \delta_p),$$

$\rho^* = \rho_{12 . 34 \ldots p}$ and

$h(F)$ denotes the pdf of non-central F with $(p-2, n-p+2)$ d. f. and non-centrality parameter γ_k^2.

Let

$$b_1(k) = c_k (\delta_1 - \underline{\beta}_2' \underline{\delta}_3) - c_k (\delta_2 - \underline{\beta}_3' \underline{\delta}_3) . \rho^* .$$

$$b_2(k) = c_k (\delta_2 - \underline{\beta}_3' \underline{\delta}_3) - c_k (\delta_1 - \underline{\beta}_2' \delta_3) . \rho^* \text{ and}$$

$$b_k^2 = (c_k^2 (\delta_1 - \underline{\beta}_2' \delta_3)^2 + c_k^2 (\delta_2 - \underline{\beta}_3' \underline{\delta}_3)^2$$

$$- 2 \rho^* c_k^2 (\delta_1 - \rho^* \underline{\beta}_3' \underline{\delta}_3) . (\delta_2 - \rho^* \underline{\beta}_2' \underline{\delta}_3)) / (1 - \rho^*)^2 ;$$

with a little algebra (2.17) reduces to

$$f_k(x) = \exp(-(b_k^2 + \gamma_k^2)/2)$$

$$\sum_{h, m, t, l=0}^{\infty} \frac{(\gamma_k^2/2)^h}{h!} \frac{(b_k^2/2)^l}{l!} \frac{(b_2^2(k)/2)^m}{m!} \frac{(2\rho^*)^t}{t!}$$

$$. \frac{(1-\rho^{*2})^{\frac{N-p+1}{2}}}{\sqrt{\pi}\, \Gamma((N-p)/2)\, \Gamma(\frac{N-p-1}{2} + m + 1)} .$$

$$. \sum_{i=0}^{m} \binom{m}{i} (2b_1(k)/b_2(k))^i \, x^{i+t} (1-x^2)^{\frac{N-p-2}{2}}$$

$$\cdot \sum_{j=0}^{m-i} \binom{m-i}{j} (b_1(k) / b_2(k))^{2j} \, \Gamma\left(\frac{N-p+1+i+t+2j}{2}\right) \cdot$$

$$\Gamma\left(\frac{N-p+1+2m+t-i-2j}{2}\right)$$

$$\cdot \; \beta\left(\frac{p-2}{2}+h+l, \; \frac{N-p+1}{2}+m\right) / \beta\left(\frac{p-2}{2}+h, \; \frac{N-p+1}{2}\right). \tag{2.18}$$

Thus the pdf of x is given by

$$h(x / \rho^*) = \sum_{k=0}^{N} \binom{N}{k} \lambda^k (1-\lambda)^{N-k} f_k(x / \rho^*) \tag{2.19}$$

Hence we get the following.

Theorem 2.

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ be independently distributed with common pdf $f(\mathbf{x})$ given by

$$f(\mathbf{x}) = \lambda \phi(\mathbf{x}; 0, \Sigma) + (1-\lambda) \phi(\mathbf{x}; \underline{\delta}, \Sigma), \quad 0 \leq \lambda \leq 1$$

where $\phi(\mathbf{x}; \underline{\mu}, \Sigma)$ denotes the p-variate normal pdf with mean vector $\underline{\mu}$ and covariance matrix Σ. Then the distribution of $x = r_{12.34 \ldots, p}$ is given by (2.19).

2.1. Special Cases

(*a*) If $\underset{-}{\Sigma} = \begin{bmatrix} 1 & 0 & \underline{0} \\ 0 & 1 & \underline{0} \\ \underline{0} & \underline{0} & I_{p-2} \end{bmatrix}$ that is, $\underline{\beta}_2 = \underline{\beta}_3 = \underline{0}, \beta_1 = 0,$

then the null distribution of $x = r_{12.34 \ldots, p}$ can easily be written from (2.19) as

$$h(x \mid \rho^* = 0) = \sum_{k=0}^{N} \binom{N}{k} \lambda^k (1-\lambda)^{N-k} f_k(x \mid \rho^* = o) \quad (2.20)$$

where

$$f_k(x \mid \rho^* = 0) = \exp(-(b_k^2 + \gamma_k^2)/2)$$

$$\sum_{h,m,l=0}^{\infty} \frac{(\gamma_k^2/2)^h}{h!} \frac{(b_k^2/2)^l}{l!} \frac{(b_2^2(k)/2)^m}{m!}$$

$$\cdot \frac{\beta(\frac{p-2}{2} + h + l, \frac{N-p+1}{2} + m)}{\beta(\frac{p-2}{2} + h, \frac{N-p+1}{2})}$$

$$\cdot \frac{1}{\sqrt{\pi}\,\Gamma(\frac{N-p}{2})\,\Gamma(\frac{N-p-1}{2} + m + 1)} \cdot$$

$$\cdot \sum_{i=0}^{m} \binom{m}{i} (2b_1(k)/b_2(k))^i \cdot x^i (1-x^2)^{\frac{N-p-2}{2}} \cdot$$

$$\cdot \sum_{j=0}^{m-i} \binom{m-i}{j} (b_1(k)/b_2(k))^{2j} \Gamma(\frac{N-p+1+i+2j}{2}) \cdot$$

$$\cdot \Gamma(\frac{N-p+1+2m-i-2j}{2}) \quad (2.21)$$

Here,

$$b_1(k) = c_k \Delta_1$$

$$b_2(k) = c_2 \Delta_2$$

$$b_k^2 = b_1^2(k) + b_2^2(k)$$

$$\gamma_k^2 = c_k^2 \Delta_3^2$$

where Δ 's are defined as

$$\Delta^2 = (\underset{-}{\mu}_1 - \underset{-}{\mu}_2)' \Sigma^{-1} (\underset{-}{\mu}_1 - \underset{-}{\mu}_2)$$

$$= \underset{-}{\delta}' \Sigma^{-1} \underset{-}{\delta}$$

$$= (\delta_1, \delta_2, \underset{-}{\delta}_3') \Sigma^{-1} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \underset{-}{\delta}_3 \end{bmatrix}$$

$$= \Delta_1^2 + \Delta_2^2 + \Delta_3^2 \text{, that is,}$$

Δ_1, Δ_2 and Δ_3 are the square root of standardized distances for the first, second and the remaining components (3, 4, ... , p).

(*b*) If $\delta_1 - \underset{-}{\beta}_2' \underset{-}{\delta}_3 = \delta_2 - \underset{-}{\beta}_3' \underset{-}{\delta}_3 = 0$, then (2.19) reduces to

$$g(x \mid \rho^*) = \frac{2^{\frac{1}{2}(N-p-1)} (1-x^{*2})^{\frac{1}{2}(N-p-1)}}{\pi \, \Gamma(N-p)} (1-x^2)^{\frac{1}{2}(N-p-2)}$$

$$\cdot \sum_{t=0}^{\infty} \frac{(2\rho^* x)^t}{t!} \cdot \Gamma^2((N-p+1+t)/2), \qquad (2.22)$$

that is the distribution of the partial correlation coefficient under the normal theory.

3. Multiple Correlation

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ be a random sample of size N on a p-random vector $\mathbf{x} = (x_1, \ldots, x_p)$ with pdf given by (1.1). We partition S as

$$S = \begin{bmatrix} s_{11} & \mathbf{s}_{12}' \\ \mathbf{s}_{12} & \mathbf{S}_{22} \end{bmatrix} \tag{3.1}$$

where $s_{11} : 1 \times 1$, $\mathbf{S}_{12} : (p-1) \times 1$ and $\mathbf{S}_{22} : (p-1) \times (p-1)$, the square of sample multiple correlation R^2 between $\mathbf{x}_1$ and $\mathbf{x}_2 = (x_2, \ldots, x_p)$ is given by

$$R^2 = \mathbf{s}_{12}' \mathbf{S}_{22}^{-1} \mathbf{s}_{12} / s_{11} \tag{3.2}$$

The population multiple correlation coefficient, ρ, between $\mathbf{x}_1$ and $\mathbf{x}_2$ is given by

$$\rho^2 = \underline{\sigma}_{12}' \Sigma_{22}^{-1} \underline{\sigma}_{12} / \sigma_{11} \tag{3.3}$$

where Σ is partitioned, as in S, that is,

$$\Sigma = \begin{bmatrix} \sigma_{11} & \underline{\sigma}_{21}' \\ \underline{\sigma}_{12} & \Sigma_{22} \end{bmatrix} \tag{3.4}$$

However, if $\mathbf{x} \sim N_p(\underline{\mu}, \Sigma)$, then it is well known, see for example Srivastava and Khatri [1979, *p.* 90] that the distribution of R^2 is given by

$$P(R^2 \mid \rho^2) = \frac{(1-\rho^2)^{\frac{N-1}{2}}}{\beta\left(\frac{p-1}{2}, \frac{N-p}{2}\right)} (R^2)^{\frac{p-3}{2}} (1-R^2)^{(N-p-2)/2}$$

$$\cdot {}_2F_1\left(\frac{N-1}{2}, \frac{N-1}{2}, \frac{p-1}{2}, \rho^2 R^2\right) \qquad (3.5)$$

where ${}_2F_1$ is the hypergeometric function. When $\rho^2 = 0$, this reduces to

$$f(R^2/\rho^2 = 0) = \left[\beta\left(\frac{p-1}{2}, \frac{N-1}{2}\right)\right]^{-1} \left(R^3\right)^{\frac{p-3}{2}} \left(1-R^2\right)^{\frac{N-p-2}{2}} \qquad (3.6)$$

If, however, the observations are not normal, the distribution of R^2 may not be as given above and the significance level and power of this test may be distorted. We now derive the null distribution of R^2 when the sample is from a mixture of two normals given in (1.1). The multiple correlation coefficient is invariant under the group G of transformation g, given by

$$gX = (x_1 + a_1, \underline{x}_2 + \underline{a}_2) \begin{bmatrix} g_{11} & \underline{0}' \\ \underline{0} & G_{22} \end{bmatrix} \qquad (3.7)$$

where $g_{11} \neq 0$, $G_{22} : (p-1) \times (p-1)$ is a nonsingular matrix and $a_1 : 1 \times 1$, $a_2 : (p-1) \times 1$ are real vectors.

Thus we assume without any loss of generality that $\underline{\mu}_1 = \underline{0}$, $\underline{\mu}_2 = \underline{\delta}$ and

$$\Sigma = \begin{bmatrix} 1 & \underline{\beta}' \\ \underline{\beta} & 1 \end{bmatrix} \qquad (3.8)$$

As usual, we shall derive the distribution of

$$w = \frac{n-p+1}{p-1} \left(R^2 / (1-R^2)\right)$$

$$= \frac{t}{s} \left(S_{12}' S_{22}^{-1} S_{12} / (s_{11} - S_{12}' S_{22}^{-1} S_{12}) \right) \qquad (3.9)$$

where $n = N-1$, $t = n-p+1$ and $s = (p-1)$. From (2.7) we can write S as

$$S = Y'Y$$

$$= \begin{bmatrix} y_1' \\ \\ Y_2' \end{bmatrix} (y_1 , Y_2) \qquad (3.10)$$

where Y is a $p \times n$ matrix with $n=N-1$, $y_1 : n \times 1$ and $Y_2 : n \times (p-1)$.

The mean of the first oolumn random vector is $c_k \underline{\delta}$ with $c_k = (k\,(N-k) / N)^{1/2}$ and the means of the remaining $(n-1)$ column random vectors are zero. In terms of Y's

$$R^2 = [\, y_1' \, Y_2 \, (Y_2' \, Y_2)^{-1} \, Y_2' \, y_1 \,] / y_1' \, y_1 \text{ , and}$$

$$w = \frac{t}{s} \, y_1' \left(\frac{Y_2 \, (Y_2' \, Y_2)^{-1} \, Y_2'}{\sigma_{1.2}} \right) y_1 \Big/ y_1' \, .$$

$$\cdot \left(\frac{I - Y_2 \, (Y_2' \, Y_2)^{-1} \, Y_2'}{\sigma_{1.2}} \right) y_1 \qquad (3.11)$$

Let $Y_2' = (\underline{\mu}_1, \underline{\mu}_2, \dots, \underline{\mu}_n)$ and $\underline{\delta}' = (\delta_1 , \delta_2')$ where $\underline{\mu}_i$'s and δ_2 are $(p-1)$ vectors. Note that

$$E\,(y_1') = c_k \, \delta_1 \, (1, 0, \dots , 0), \; E\,(\underline{u}_1) = c_k \, \underline{\delta}_2$$

and

$$E(\mathbf{u}_i) = \underline{0},\ i = 2, 3, \dots, n.$$

Given Y_2, $\mathbf{y}_1$ has a normal distribution with mean vector $c_k(\delta_1 - \underline{\beta}'\underline{\delta}_2)(1, 0, \dots, 0) + \underline{\beta}' Y_2'$ and covariance matrix $\sigma_{1.2}\,\mathrm{I}$ $(\sigma_{1.2} = 1 - \underline{\beta}'\underline{\beta},\ \underline{\beta}'\underline{\beta} = \rho^2)$. Hence the denominator of w for fixed k and given Y_2 has a noncentral chi-square distribution with t d. f. and noncentrality parameter

$$\eta_k^2 = c_k^2 \frac{(\delta_1 - \underline{\beta}'\delta)^2}{\sigma_{1.2}} (1, 0, \dots, 0)\,[\mathrm{I} - Y_2 (Y_2' Y_2)^{-1} Y_2]\,(1, 0, \dots, 0),$$

Let $\tilde{\delta}_k^2 = c_k^2 (\delta_1 - \underline{\beta}'\delta_2)^2 / \sigma_{1.2}$ and $B = \sum_{i=2}^{n} \mathbf{u}_i \mathbf{u}_i'$. Then

$$B \sim W_s(\mathrm{I}, n-1) \text{ and } Y_2' Y_2 = \mathbf{u}_1 \mathbf{u}_1' + B.$$

Hence,

$$\eta_{2k}^2 = \tilde{\delta}_k^2 (1 + \mathbf{u}_1' B^{-1} \mathbf{u}_1) \tag{3.12}$$

Similarly, given Y_2, the numerator has also a non-central chi-square distribution with s d. f. and noncentrality parameter

$$\eta_{1k}^2 = \tilde{\delta}_k^2 \frac{\mathbf{u}_1 B^{-1} \mathbf{u}_1}{1 + \mathbf{u}_1' B^{-1} \mathbf{u}_1} \tag{3.13}$$

where

$$\gamma^2 = \beta' Y_2' Y_2 \underline{\beta} / \sigma_{1.2}\,. \tag{3.14}$$

But $\sigma_{1.2}\gamma^2 / \underline{\beta}'\underline{\beta}$ has a non-central chi-square distribution with n d. f. and noncentrality parameter

$$\xi_k^2 = c_k^2 (\underline{\beta}' \underline{\delta}_2)^2 / \underline{\beta}' \underline{\beta}, \tag{3.15}$$

and

$$v = \frac{t}{s} \mathbf{u}_1' \mathbf{B}^{-1} \mathbf{u}_1 \sim F_{s,t}(\gamma_k^2) \tag{3.16}$$

where $F_{s,t}(\gamma_k^2)$ denotes a non-central F with (s, t) d. f. and noncentrality parameter γ_k^2 given by

$$\gamma_k^2 = c_k^2 \underline{\delta}_2' \underline{\delta}_2 . \tag{3.17}$$

Hence, for fixed k and given v and γ^2

$$w = \frac{t}{s} \frac{R^2}{1-R^2} \sim F_{s,t}(\eta_{1k}^2, \eta_{2k}^2) \tag{3.18}$$

Here, $F_{s,t}(\eta_{1k}^2, \eta_{2k}^2)$ denotes the doubly non-central F-distribution with (s, t) degrees of freedom and non-centrality parameters η_{1k}^2 and η_{2k}^2.

Thus, for fixed k the pdf of w is given by

$$g_k(w) = \int_0^\infty \int_0^\infty \exp - \frac{1}{2}\left(\eta_{1k}^2 + \eta_{2k}^2 \right)$$

$$\sum_{i,j=0}^{\infty} \frac{\left(\eta_{1k}^2/2\right)^i}{i!} \frac{\left(\eta_{2k}^2/2\right)^j}{j!} \cdot \frac{s}{t}\left(\frac{s}{t} w\right)^{\frac{s}{2}+i-1}$$

$$\cdot \left(1+\frac{s}{t}w\right)^{-\left(\frac{s+t}{2}+i+j\right)} \cdot h_{1k}(v) \cdot h_{2k}(\gamma^2)\, dv\, d\gamma^2. \tag{3.19}$$

where

$$h_{1k}(v) = \exp\left(-\frac{1}{2}\gamma_k^2\right) \cdot \frac{s}{t} \times$$

$$\times \sum_{l=0}^{\infty} \frac{\left(\gamma_k^2/2\right)^l}{l\,!} \left(\frac{s}{t}v\right)^{\frac{s}{2}+l-1} \left(1+\frac{s}{t}v\right)^{-\left(\frac{s+t}{2}+l\right)} \cdot$$

$$\cdot \ \beta\left(\frac{s}{2}+l,\, t/2\right)$$

and

$$h_{2k}(\gamma^2) = \frac{\sigma_{1.2}}{2\,\underline{\beta}'\,\underline{\beta}} \exp - \frac{1}{2}\left(\xi_k^2\right) \times$$

$$\times \sum_{m=0}^{\infty} \frac{\left(\xi_k^2/2\right)^m}{m\,!} \left(\frac{\sigma_{1.2}\,\gamma^2}{2\,\underline{\beta}'\,\underline{\beta}}\right)^{\frac{n}{2}+m-1}$$

$$\exp - \frac{1}{2}\left(\frac{\sigma_{1.2}}{\underline{\beta}'\,\underline{\beta}}\right) / \Gamma\left(\frac{n}{2}+m\right)$$

After simplification, the distribution of w for fixed k is given by

$$g_k(w) = \exp - \frac{1}{2}\left(\tilde{\delta}_k^2 + \gamma_k^2 + \xi_k^2\right) \times$$

$$\times \sum_{j,\,l,\,m=0}^{\infty} \frac{(\tilde{\delta}_k^2/2)^j}{j\,!} \frac{(\gamma_k^2/2)^l}{l\,!} \frac{(\xi_k^2/2)^m}{m\,!} \times$$

$$\cdot \sum_{i=0}^{\infty} \sum_{q=0}^{i} \frac{\left(\tilde{\delta}_k^2 / 2\right)^q}{q\,!} \cdot \frac{\left(\dfrac{\underline{\beta}'\,\underline{\beta}}{\underline{\beta}'\,\underline{\beta} + \sigma_{1.2}}\right)^{i-q}}{(i-q)\,!} \times$$

$$\times \left(\frac{\sigma_{1.2}}{\underline{\beta}'\,\underline{\beta} + \sigma_{1.2}}\right)^{\frac{n}{2}+m} \frac{\Gamma\left(\frac{n}{2}+m+i-q\right)}{\Gamma\left(\frac{n}{2}+m\right)} \times$$

$$\times \frac{\beta\left(\frac{s}{2}+l+q,\ \frac{t}{2}+j\right)}{\beta\left(\frac{s}{2}+l,\ t/2\right)} \times$$

$$\times \frac{\frac{s}{t}\left(\frac{s}{t}w\right)^{\frac{s}{2}+i-1}\left(1+\frac{s}{t}w\right)^{-\frac{1}{2}(s+t+2i+2j)}}{\beta\left(\frac{s}{2}+i,\ \frac{t}{2}+j\right)},$$

$$0 < w < \infty \qquad (3.20)$$

Thus, the p.d.f. of w is given by

$$g_N(w) = \sum_{k=0}^{N} \binom{N}{k} \lambda^k (1-\lambda)^{N-k} g_k(w) \qquad (3.21)$$

Hence, we get,

Theorem 3. Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ be independentiy distributed with p.d.f. given by

$$f(\mathbf{x}) = \lambda\,\phi\,(\mathbf{x}\,;\,0,\,\Sigma) + (1-\lambda)\,\phi\,(\mathbf{x}\,;\,\underline{\delta}\,,\,\Sigma), \qquad 0 \leq \lambda \leq 1$$

where $\Sigma = \begin{bmatrix} 1 & \underline{\beta}' \\ \underline{\beta} & \mathbf{I} \end{bmatrix}$. Then the distrlbution of

$w = \frac{s}{t} R^2 (1-R^2)^{-1}$ where $s = p-1$, $t = n-p+1$ and $n = N-1$ is given by (3.21).

3.1 Special Cases.

(*a*) Under the null hypothesis, $\rho^2 = 0$ the p.d.f. of w given by (3.21) reduces to

$$g(w) = \sum_{k=0}^{N} \binom{N}{k} \lambda^k (1-\lambda)^{N-k} \exp - \frac{1}{2}\left(c_k^2 \delta_1^2 + c_k^2 \underset{-}{\delta}_2' \underset{-}{\delta}_2 \right)$$

$$\cdot \sum_{i,j,l=0} \frac{(c_k^2 \delta_1^2 / 2)^{i+j}}{i!\,j!} \frac{(c_k^2 \underset{-}{\delta}_2' \underset{-}{\delta}_2 / 2)^l}{l!}$$

$$\cdot \frac{\beta\left(\frac{s}{2} + l + i, \frac{t}{2} + j\right)}{\beta\left(\frac{s}{2} + l, t/2\right)} .$$

$$\cdot \frac{\frac{s}{t}\left(\frac{s}{t} w\right)^{\frac{s}{2}+i-1} \left(1+\frac{s}{t} w\right)^{-\left(\frac{s+t}{2}+i+j\right)}}{\beta\left(\frac{s}{2}+i, \frac{t}{2}+j\right)}$$

$$0 \leq w \leq \infty .$$

This null distribution has been derived by Srivastava (1981) and has been used by Awan (1981) to study the robustness of R^2 test.

(*b*) If $\delta_1 - \underset{-}{\beta}' \underset{-}{\delta}_2 = 0$, then (3.21) reduces to

$$g(w) = \sum_{k=0}^{N} \binom{N}{k} \lambda^k (1-\lambda)^{N-k} . \exp - \frac{1}{2} \xi_k^2$$

$$\sum_{m,i=0}^{\infty} \frac{(\xi_k^2/2)^m}{m!} \frac{\left(\frac{\underset{-}{\beta}' \underset{-}{\beta}}{\underset{-}{\beta}' \underset{-}{\beta} + \sigma_{1.2}}\right)^i}{i!} \left(\frac{\sigma_{1.2}}{\sigma_{1.2} + \underset{-}{\beta}' \underset{-}{\beta}}\right)^{\frac{n}{2}+m}$$

$$\cdot \frac{\left(\frac{n}{2}+m+i\right)}{\Gamma\left(\frac{n}{2}+m\right)} \cdot$$

$$\cdot \frac{s}{t}\left(\frac{s}{t}w\right)^{-\frac{s}{2}+i-1}\left(1+\frac{s}{t}w\right)^{-\left(\frac{s+t}{2}+i\right)} \Big/$$

$$\beta\left(\frac{s}{2}+i, \frac{t}{2}\right).$$

This special case has been used by Awan [1981] to study the behaviour of power of R^2 – test.

REFERENCES

1. Anderson, T.W. (1958). An Introduction to Multivariate Statistical Analysis, Wiley, New York.
2. Awan, H.M. and Srivastava, M.S. (1980). "Inference robustness studies of the location parameter of a class of contaminated normal distribution". Technical Report ‡ 8, Dept. of Statistics, University of Toronto.
3. Awan, H.M. (1981). On the robustness of T^2, r, $r_{12.34 \cdots p}$ and R^2 when sampling from a contaminated normal distribution. Unpublished Ph. D. Thesis, Dept. of Statistics, University of Toronto.
4. Baker, G.A. (1932). Distribution of the means divided by the standard deviations of samples from non-homogeneous populations. Ann. Math. Statist. **3**, 1-9.
5. Fisher, R.A. (1915). The frequency distribution of values of the correlation coefficient in samples from an indefinitely large population. Biometrika **10**, 507-21.
6. ——————— (1924). The distribution of the partial correlation coefficient. Metron **3**, 329-32.

7. ———————— (1928). The general sampling distribution of the multiple correlation coefficient. Proc. Roy. Soc. A, **121**, 654-73.

8. Srivastava, M.S. and Khatri, C.G. (1979). An introduction to multivariate statistics. North-Holland, New York.

9. Srivastava, M.S. and Awan, H.M. (1980). On the robustness of correlation coefficient in sampling from a mixture of two bivariate normals. Tech. Report No. 9, Dept. of Statistics, University of Toronto.

10. Srivastava, M.S. (1981). On the distribution of Hotelling's T^2 and multiple correlation R^2 when sampling from a mixture of two normals. Tech. Report, Dept. of Statistics, University of Toronto.

11. Srivastava, M.S. and Awan, H.M. (1982). On the robustness of Hotelling's T^2—test and distributions of linear and quadratic forms in sampling from a mixture of two multivariate normal populations. Comm. in Statist.

The Punjab University
Journal of Mathematics
Vol. XIV—XV (1981-82), pp. 109-116

A CLASS OF ESTIMATORS OF RATIO (PRODUCT) IN SAMPLE SURVEYS

By

R. KARAN SINGH AND GURDEEP SINGH
Lucknow University, Lucknow, INDIA.

Summary : For the estimation of ratio (product) of two population means, a class of estimators is proposed, using auxiliary information on two auxiliary variables, of which the estimators by Singh (1965, 67, 69), Rao and Pareira (1968) and Shah and Shah [978] are particular cases. The mean squars error of the proposed estimator is found to the first degree of approximation and a comparative study is made among various estimators.

Introduction.

Let y_1, y_2 be the variables of our interest with population means $\bar{Y}_1$, $\bar{Y}_2$ and x_1 and x_2 be the auxiliary variables with population means $\bar{X}_1$, $\bar{X}_2$ respectively. Also, let $R = \bar{Y}_1 / \bar{Y}_2$ and $P = \bar{Y}_1 . \bar{Y}_2$ be respectively the ratio and product of the population means $\bar{Y}_1$ and $\bar{Y}_2$.

The usual estimators of R and P are $r = \bar{y}_1 / \bar{y}_2$ and $p = \bar{y}_1 . \bar{y}_2$ where $\bar{y}_1$ and $\bar{y}_2$ are the unbiased estimators of $\bar{y}_1$ and $\bar{y}_2$. Let $\bar{x}_1$ and $\bar{x}_2$ be the unbiased estimators of $\bar{X}_1$ and $\bar{X}_2$; S_{12} be the population covariance between x_1 and x_2 with s_{12} being the unbiased estimator of S_{12}.

When the population means $\bar{X}_1$ and $\bar{X}_2$ of x_1 and x_2 are known, the estimators proposed by Singh [1969] Rao and Pareira [1968] and Shah and Shah [1978] for the estimation of R (or P) are given by

$$r_1 = r\,(\bar{x}_1 / \bar{X}_1)^{\alpha_1}\,(\bar{x}_2 / \bar{X}_2)^{\alpha_2},$$

$$r_2 = r\left[W_1 (\bar{x}_1 / \bar{X}_1)^{\alpha_1} + W_2 (\bar{x}_2 / \bar{X}_2)^{\alpha_2} \right]$$

$$r_3 = r\left[W_1 (\bar{x}_1 / \bar{X}_1)^{\alpha_1 / W_1} + W_2 (\bar{x}_2 / \bar{X}_2)^{\alpha_2 / W_2} \right]$$

$$r_4 = r\left[W_1 (\bar{x}_1 / \bar{X}_1) + W_2 (\bar{x}_2 / \bar{X}_2) \right]^{\alpha}$$

where W_1 and W_2 are weights such that $W_1 + W_2 = 1$ and α's are constants to be determined by minimizing the mean square error of the corresponding estimator.

Similar estimators of the product $P = \bar{Y}_1 . \bar{Y}_2$ can be written by replacing r by p in r_i , $i = 1,2,3,4$. All the estimators r_i , i=1,2,3,4 may be identified as the particular cases of the generalized estimator

$$t_1 = r\,f(u, v)$$

where $u = \bar{x}_1 / \bar{X}_1$, $v = \bar{x}_2 / \bar{X}_2$; (u, v) assumes values in a bounded, closed, convex subset I of the two dimensional real space containing the point (1,1), f is a bounded and continuous function of u and v with continuous and bounded first and second partial derivatives in I and $f(1, 1) = 1$.

Assuming that the covariance S_{12} between two auxiliary variables x_1 and x_2 is known, an improved estimator is suggested in this paper which is more efficient than the estimator t_1 and hence than r_i , $i = 1, 2, 3, 4$.

Denoting $w = s_{12} / S_{12}$, let (u, v, w) assume values in a bounded, closed, convex subset I′ of the three dimensional real space containing the point (1, 1, 1). Let $f(u, v, w)$ such that $f(1, 1, 1) = 1$, be a bounded and continuous function of u, v and w with bounded first and second partial derivatives in I′. Then the proposed generalized estimator is given by

$$t_2 = r\, f(u, v, w) \tag{1.2}$$

2. Mean Square Error (MSE) of Estimators of R.

Let us denote by

C_{0i} = Coefficient of variation of $\bar{y}_i$; $i = 1, 2$

C_j = Coefficient of variation of $\bar{x}_j$; $j = 1, 2$

C_3 = Coefficient of variation of s_{12}

ρ = Correlation coefficient between $\bar{y}_1$ and $\bar{y}_2$

ρ_{12} = Correlation coefficient between $\bar{x}_1$ and $\bar{x}_2$

ρ_{0ij} = Correlation coefficient between $\bar{y}_i$ and $\bar{x}_j$; $i, j = 1, 2$

ρ_{0i3} = Correlation coefficient between $\bar{y}_i$ and s_{12} ; $i = 1, 2$

ρ_{i3} = Correlation ceefficient between $\bar{x}_i$ and s_{12} ; $i = 1, 2$

$C_{ij} = \rho_{ij}\, C_i\, C_j$; $i = 1, 2$; $j = 1, 2, 3$

$C_{0ij} = \rho_{0ij}\, C_{0i}\, C_j$, $i = 1, 2$, $j = 1, 2, 3$

$d_k = C_{01k} - C_{02k}$; $k = 1, 2, 3$

Also, let $A_1 = d_1 C_2^2 - d_2 C_{12}$, $A_2 = d_2 C_1^2 - d_1 C_{12}$,

$A = A_1 d_1 + A_2 d_2$ and $B = C_1^2 C_2^2 - C_{12}^2$.

Let $\bar{y}_1 = \bar{Y}_1 (1 + e_3)$, $\bar{y}_2 = \bar{y}_2 (1 + e'_0)$, $\bar{x}_1 = \bar{X}_1 (1 + e_1)$, $\bar{x}_2 = \bar{X}_2 (1 + e_2)$, $s_{12} = S_{12} (1 + e_3)$ where it is assumed that the sample is large enough to make $| e'_0 |$ and $| e_i |$, $i = 0, 1, 2, 3$ so small that the terms of degree greater than two in e'_i s may be neglected to justify the first degree approximation.

Now

MSE $(t_1) = E(t_1 - R)^2$ from which, to the first degree of approximation.

$$\text{MSE}(t_1) = R^2 E\left(e_0 - e_0' + e_1 f_u'(1,1) + e_2 f'(1,1) \right)^2$$

$$= \text{MSE}(r) + R^2 \Big(C_1^2 \left\{ f_u'(1,1) \right\}^2 + C_2^2 \left\{ f_v' \; 1,1) \right\}^2$$

$$+ 2C_{12} f_u'(1,1) f_v'(1,1) + 2d_1 f_u'(1,1) + 2d_2 f_v'(1,1) \Big) \quad (2.1)$$

where $\text{MSE}(r) = R^2 \left[C_{01}^2 + C_{02}^2 - 2\rho\, C_{01}\, C_{02} \right]$

$f_u'(1,1)$ and $f_v'(1,1)$ are the first partial derivatives of $f(u, v)$ with respect to u and v at the point (1, 1).

From (2.1) the mean square errors of the estimators r_i, $i=1, 3, 4$ can be easily written. For $f(u, v) = u^{\alpha_1} v^{\alpha_2}$, $(w_1 u^{\alpha_1} + w_2 v^{\alpha_2})$, $(w_1 u^{\alpha_1/w_1} + w_2 v^{\alpha_2/w_2})$, $(w_1 u + w_2 v)^{\alpha}$ the mean square errors of the estimators r_i, $i = 1, 2, 3, 4$ are given by

$$\text{MSE}(r_1) = \text{MSE}(r) - R^2 (C_1^2 \alpha_1^2 + C_2^2 \alpha_2^2 + 2C_{12}\alpha_1\alpha_2 + 2d_1\alpha_1 + 2d_2\alpha_2)$$

$$\text{MSE}(r_2) = \text{MSE}(r) - R^2 (C_1^2 \alpha_1^2 w_1^2 + C_2^2 \alpha_2^2 w_2^2 + 2\alpha_1\alpha_2 w_1 w_2 C_1 + 2d_1\alpha_1 w_1 + 2d_2\alpha_2 w_2)$$

$$\text{MSE}(r_3) = \text{MSE}(r) = R^2 (C_1^2 \alpha_1^2 + C_2^2 \alpha_2^2 + 2C_{12}\alpha_1\alpha_2 + 2d_1\alpha_1 + 2d_2\alpha_2)$$

$$\text{MSE}(r_4) = \text{MSE}(r) - R^2 (C_1^2 \alpha_1 w_1^2 + C_2^2 \alpha_2^2 w_2^2 + 2C_{12}\alpha_2 w_1 w_1 + 2d_1\alpha w_1 + 2\, d_2 \alpha w_2)$$

The optimum choice of the function $f(u, v)$ minimizing (2.1) is f such that

$$f_u'(1,1) = -A_1/B \text{ and } f_v'(1,1) = -A_2/B$$

and

$$\text{min. MSE}(t_1) = \text{MSE}(r) - R^2 A/B \tag{2.2}$$

Now expanding $f(u, v, w)$ in a second order Taylor series about the point (1, 1, 1), we have

$$t_1 = R(1 + e_0)(1 + e_0')^{-1}\Big[1 + e_1 f_u'(1,1,1) + e_2 f_v'(1,1,1) + e_3 f_w'(1,1,1) + \tfrac{1}{2}\Big(\frac{e_1 \partial}{\partial u} + \frac{e_2 \partial}{\partial v} + \frac{v_2 \partial}{\partial w}\Big)^2 f(u^*, v^*, w^*)\Big]$$

where $u^* = 1 + Be_1$, $v^* = 1 + Be_2$, $w^* = 1 + Be_3$, $0 < B < 1$ and $f'(1, 1, 1)$ denote the first partial derivatives of $f(u, v, w)$ with respect to the variable denoted in suffix at the point (1, 1, 1).

Upto the first degree of approximation,

$$\begin{aligned}\text{MSE}(t_2) &= E(t_2 - R)^2 \\ &= R^2 E\Big[e_0 - e_0' + e_1 f_u'(1,1,1) + e_2 f_v'(1,1,1) + e_3 f_w'(1,1,1)\Big]^2 \\ &= \text{MSE}(r) + R^2\Big[C_1^2\{f_u'(1,1,1)\}^2 + C_2^2\{f_v'(1,1,1)\}^2 \\ &\quad + C_3^2\{f_w'(1,1,1)\}^2 + 2(C_{12} f_u'(1,1,1) f_v'(1,1,1) \\ &\quad + C_{13} f_u'(1,1,1) f_w'(1,1,1) + C_{23} f_v'(1,1,1) f_w'(1,1,1) \\ &\quad + d_1 f_u'(1,1,1) + d_2 f_v'(1,1,1) + d_3 f_w'(1,1,1))\Big]\end{aligned} \tag{2.3}$$

The optimum choice of the function $f(u, v, w)$ minimizing MSE (t_2) is f such that

$$f_u'(1,1,1) = -\frac{1}{S}\Big[d_1 (C_2^2 C_3^2 - C_{23}^2) - d_2 (C_3^2 C_{12} - C_{13} C_{23}) - d_3 (C_2^2 C_{13} - C_{12} C_{23}) \Big]$$

$$f_v'(1,1,1) = -\frac{1}{S}\Big[d_2 (C_1^2 C_3^2 - C_{13}^2) - d_1 (C_3^2 C_{12} - C_{13} C_{23}) - d_3 (C_1^2 C_{23} - C_{12} C_{13}) \Big] \qquad (2.4)$$

$$f_w'(1,1,1) = -\frac{1}{S}\Big[d_3 (C_1^2 C_2^2 - C_{12}^2) - d_1 (C_2^2 C_{13} - C_{12} C_{23}) - d_2 (C_1^2 C_{23} - C_{13} C_{12}) \Big]$$

where $S = C_1^2 (C_2^2 C_3^2 - C_{23}^2) - C_{12} (C_{12} C_3^2 - C_{13} C_{23}) + C_{13} (C_{12} C_{23} - C_{13} C_2^2)$

and

$$\text{min. MSE}(t_2) = \text{min. MSE}(t_1) - R^2\, T / T' \qquad (2.5)$$

with $T = \Big\{ (\rho_{013} - \rho_{023})(1-\rho_{12}^2) + \rho_{12} (\rho_{23} \rho_{013} - \rho_{23} \rho_{023} + \rho_{13} \rho_{012} - \rho_{13} \rho_{022}) - (\rho_{011} - \rho_{021}) \rho_{13} - (\rho_{012} - \rho_{022}) \rho_{23} \Big\}^2 \geq 0$

and

$$T' = (1-\rho_{12}^2)^2 \left\{ \frac{S}{C_1^2 C_2^2 - C_{12}^2} \right\} \geq 0$$

implying that

$$\text{min. MSE } (t_2) < \text{min. MSE } (t_1) \tag{2.6}$$

From (2.6) it is obvious that t_2 is more efficient than t_1 and hence than r_i , $i = 1, 2, 3, 4$.

Some members of the class of estimators represented by t_2 are

$$t_{2(1)} = ru^{\alpha} v^{\beta} w^{\delta}, \quad t_{2(2)} = r\left[1+\alpha(u-1)+\beta(v-1)+\delta(w-1)\right]$$

$$t_{2(3)} = r\left[1-\alpha(u-1)-\beta(v-1)-\delta(w-1)\right]^{-1} \text{ etc.}$$

If α , β and δ in above estimators are respectively given by the right hand sides of the equations in (2.4), $t_{2(i)}$, $i = 1, 2, 3$ attain the lower bound of the variance given by (2.5).

Next, if we consider a wider class of estimators $t_3 = g(r, u, v, w)$ of R, where $g(R, u, v, w) = R$ and $g'(R. u, v, w) = 1$, g' being the first partial derivative of g with respect to r, the minimum mean square error of t_3 is the same as that of t_2 given by (2.5) and is not reduced. Also, the regression type estimator $r + \alpha(u-1) + \beta(v-1) + \delta(w-1)$ is a member of the class represented by t_3 but not of t_2 . Thus, the minimum mean square error of the regression type estimator is attained by estimators from the class t_2 .

Similar estimators for the product P can be defined and similar results can be obtained.

ACKNOWLEDGEMENT

The authors wish to thank the refree for his comments. This work is partialy supported by CSIR, India.

REFERENCES

1. Murthy, M.N. (1967), SAMPLING THEORY AND METHODS, Statistical Publishing Society, India.

2. Rao, J.N.K. and Pereira, N.P. (1968) – On double ratio estimators Sankhyā, **A-30**, 83-90.

3. Shah, S.M., and Shah, D.N. (1978) – Ratio-cum-product estimators for estimating ratio (product) of population parameters, Sankhyā **C-40**, 155-166.

4. Singh, M.P. (1965), On the estimators of ratio and product of the population parameters, Sankhyā, **B-27**, 321-328.

5. Singh, M.P. (1267), Ratio-cum-product mathod of estimation Matrika, **12** (**1**), 34-43.

6. Singh, M.P. (1969), Comparison of some ratio-cum-product estimators, Sankhyā **B-31**, 375-378.

The Punjab University
Journal of Mathematics
Vol. XIV—XV (1981-82), pp. 117-122

ON THE GRIEGO-HERSH APPROACH TO RANDOM EVOLUTIONS

By

MANUEL KEEPLER
Departmeut of Mathematics
and Computer Science
South Carolina State College
Orangeburg, South Ca.olina 29117
U.S.A.

Griego and Hersh [1, 2] introduced the random evolution of a family of semigroups with switching among semigroups controlled by a finite Markov chain to provide a probabilistic approach to the study of a class of Cauchy problems. In [3] the author showed that the random evolutions of Griego and Hersh were backward random evolutions and constructed forward random evolutions. Forward random evolutions provide a probabilistic approach to the study of a different but analogous class of Cauchy problems. The proofs in [3-5] utilize the method of "renewal equations" in Markov chains. However, it would be of interest to directly obtain forward random evolutions from the Markov chain. This would give additional insight as to how the random evolution structure relates to the mechanism of the of the Markov chain. In this note we carry out this formulation, which turns out to be simpler than the formulation in [3]. In a forthcoming paper [6], the two formulations are combined to present random evolutions as Markov processes.

Suppose $v = [\, v(t),\ t \geq 0 \,]$ is a Markov chain with state space (1......N), stationery transition probabilities $p_{ij}(t)$, and infinitesimal martix $Q = < q_{ij} > = < p'_{ij}(0) >$. P_i is the probability measure defined on sample paths $w(t)$ for v under the condition $w(0) = i$. E_i

denotes integration with respect to P_i. For a sample path $\omega \varepsilon \Omega$ of v, $\tau_j(\omega)$ is the time of the j_{th} jump, and $N(t, \omega)$ is the number of jumps up to time t.

Let $\{T_i(t),\ t \geq 0,\ i = 1 \ldots, N\}$ be a family of strongly continuous semigroups of boundnd linear operators on a fixed Banach space B. A_i is the infinitesimal generator of T_i and D_i is the domain of A_i.

Definition 1. A forward random evolution $\{S(t, \omega),\ t \geq 0\}$ is defined by the product

$$S(t) = T'_{v(\tau_{N(t)})}(t-\tau_{N(t)}) \ldots T_{v(\tau_1)}(\tau_2 - \tau_1)\, T_{v(0)}(\tau_1).$$

The proof of the following lemma is parallel to that of Lemma 2 in [2] and is thus omitted.

Lemma. If $g : \Omega \to B$ is Bochner P_i –integrable for a fixed $i = 1, \ldots, N$, then for each $t \geq 0$ the function $\omega \to S(t, \omega)\, g(\omega)$ is Bochner P_i – integrable and

$$E_i[\, S(t)\, g \mid F_t\,](\omega) = S(t, \omega)\, E_i[\, g \mid F_t\,](\omega), \tag{2.1}$$

for almost all ω with respect to P_i, where F_t is the σ–algebra generated by the random variables $v(u)$, $0 \leq u \leq t$, that is, F_t is the past up to time t for v.

Let $\tilde{B}$ be the N-fold Cartesian product of B with itself. A typical element of $\tilde{B}$ is denoted by $\tilde{f} = \langle f_i \rangle$ where $f_i \varepsilon B$, $i = 1, \ldots, N$. We equip $\tilde{B}$ with any appropriate norm so that $\| \tilde{f} \| \to 0$ as $\| f_i \| \to 0$ for each i.

Definition 2. For $t \geq 0$ define the (matrix) operator $\tilde{S}(t)$ on $\tilde{B}$ specified componentwise by

$$(\tilde{S}(t)\tilde{f})_k = \sum_{i=1}^{N} E_i[S(t) f_i \; ; \; v(t) = k],$$

where

$$E_i[S(t) f_i \; ; \; v(t) = k] = E_i\left[\, S(t) f_i\, I_{\{v(t) = k\}} \right].$$

Theorem 1. $\{\tilde{S}(t),\ t \geq 0\}$ is a strongly continuous semigroup of bounded linear operators on $\tilde{B}$.

Proof. The fact that $\tilde{S}(t)$ is a bounded linear operator follows by an argument analogous to that used in the last part of the proof of Lemma 2 in [2[. Also, since $v(\,.\,,\omega)$ is continuous at a fixed t for almost all sample paths ω, it follows that $S(t)$ is strongly continuous in t. To complete the proof, it suffices to show that

$$(\tilde{S}(t+s)\tilde{f})_k = (\tilde{S}(t)\,\tilde{S}(s)\,\tilde{f})_k\,,$$

for each k.

Let $\theta_s\,\omega$ be defined on Ω by the requirement that $v(u,\theta_s\,\omega) = v(u+s,\omega)$, that is, θ_s shifts paths. Define $g\circ\theta_s$ by $(g\circ\theta_s)(\omega) = g(\theta_s\,\omega)$. Then the Markov property of v is expressed by the formula

$$E_i\,[g\circ\theta_s \mid F_s\,](\omega) = E_{v\,(s,\,\omega)}\,[g] \tag{2.2}$$

for almost all $\omega\,[P_i\,]$. We can omit ω and write simply $E_i\,[g\circ\theta_s \mid F_s] = E_{v\,(s)}\,[g]$.

Now, it is easy to see that $S(t)$ satisfies the equation

$$S(t+s,\omega) = S(t,\ \theta_s\,\omega)\,S(s,\omega)$$

or

$$S(t+s) = S(t)\circ\theta_s\,S(s). \tag{2.3}$$

Also,

$$I_{\{v\ (t+s)\,=\,k\}} = I_{\{\,v\,(t)\,=\,k\}}\circ\theta_s\,. \tag{2.4}$$

As a result, for fixed k, we have

$$(\tilde{S}(t+s)\,\tilde{f})_k$$

$$= \sum_{i=1}^{N} E_i\,[S(t+s)\,f_i\ ;\ v(t+s) = k]$$

$$= \sum_{i=1}^{N} E_i \, [E_i \, [S\,(t+s)\, f_i \; ; \; v\,(t+s) = k \mid F_s\,]]$$

$$= \sum_{i=1}^{N} E_i \, [E_i \, [\, S\,(t) \circ \theta_s \, I_{\{v\,(t) = k\}} \circ \theta_s \, S\,(s)\, f_i \mid F_s \,]] \qquad \text{(by 2.3 and 2.4)}$$

$$= \sum_{i=1}^{N} E_i \, [E_{v(s)} \, [\, S\,(t)\, I_{\{v\,(t) = k\}} \, E_i \, [S\,(s)\, f_i \mid F_s\,]]] \quad \text{(by 2.2)}$$

$$= \sum_{l=1}^{N} E_l \, [S\,(t)\, I_{\{v\,(t) = k\}} \sum_{i=1}^{N} E_i \, [S\,(s)\, f_i \, I_{\{v\,(s) = l\}}\,]] \qquad \text{(by 2.1)}$$

$$= (\tilde{S}\,(t)\, \tilde{S}\,(s)\, \tilde{f}\,)_k \, . \qquad \text{Q.E.D.}$$

Theorem 2. The Cauchy problem for an unknown vector $\tilde{u}\,(t)$, $t > 0$,

$$\frac{\partial u_k}{\partial t} = A_k \, u_k + \sum_{i=1}^{N} q_{ik} \, u_i \, , \quad \tilde{u}\,(0+) = \tilde{f}$$

is solved by $\tilde{u}\,(t) = \tilde{S}\,(t)\, \tilde{f}$.

Proof Let τ denote the last jump of v in the time interval $(0, t)$. Now,

$$(\tilde{S}\,(t)\, \tilde{f}\,)_k \, .$$

$$= \sum_{n=0}^{\infty} \sum_{i=1}^{N} E_i \, [S\,(t)\, f_i \; ; \; N\,(t) = n, \; v\,(t) = k]$$

$$= e^{-q_k t} \, T_k\,(t)\, f_k$$

$$+ \sum_{n=1}^{\infty} \sum_{i=1}^{N} \int_0^t E_i \, [S\,(t)\, f_i \; ; \; N\,(t) = n, \; v\,(\tau_n) = k \mid v\,(s)\,,\; s \leq \tau'\,]\; P_i\,(\tau \,\varepsilon\, d\tau')$$

$$= e^{-q_k t}\, T_k(t)\, f_k$$

$$+ \sum_{n=1}^{\infty} \sum_{i=1}^{N} \sum_{j \neq k} \int_0^t T_k(t-s)\, E_i\, [S(s)\, f_i\ ;\ N(s)$$

$$= n-1,\ v(\tau_{n-1}) = j\,]\ q_{ik}\, e^{-q_k (t-s)}\, ds$$

$$= c^{-q_k t}\, T_k(t)\, f_k + \int_0^t T_k(t-s) \sum_{j \neq k} (\tilde{S}(s)\, \tilde{f}\,)_j\ \cdot$$

$$\cdot\ q_{jk}\, e^{-q_k (t-s)}\, ds.$$

Let $f_i \varepsilon D_i$. Then,

$$\lim_{t \to 0} \frac{1}{t} [\,(\tilde{S}(t)\, \tilde{f}\,)_k - f_k\,]$$

$$= \lim_{t \to 0} \frac{1}{t} [\, e^{-q_k t}\, T_k(t)\, f_k - f_k\,]$$

$$+ \lim_{t \to 0} \frac{1}{t} [\int_0^t T_k(t-s) \sum_{j \neq k} (\tilde{S}(s)\, \tilde{f}\,)_j\ q_{jk}\, e^{-q_k (t-s)}\, ds$$

$$= A_k\, f_k + \sum_{i=1}^{N} q_{ik}\, f_i\ .$$

By standard semigroup theory we obtain the result of the theorem.

REFERENCES

1. Griego, R.J. and R. Hersh, Theory of random evolutions with applications to partial differential equations, *Trans. Amer. Math. Soc.*, **156** (1971), 405-418.

2. Hersh, R., Random evolutions : a survey of results and problems, *Rocky Mountain Math. J.*, **4** (1974), 443-477.

3. Keepler, M., Backward and forward random evolutions, *Indiana Univ. Math. J.*, **24** (1975), 937-947.

4. ————, Perturbation theory for backward and forward random evolutions, *J. Math, Kyoto Univ.*, **16** (1976), 395-411.

5. ————, On random evolutions induced by countable state space Markov chains, *Portugaliae Math.*, **37** (1981), 203-207.

6. ————, Random evolutions are Markov processes, *Bulletin Inst. Math, Stat.* **5** (1 76), Abstract 76 to 91.

The Punjab University
Journal of Mathematics
Vol. XIV–XV (1981-82), pp. 123-128

A NOTE ON THE CONVERGENCE OF SEQUENCE OF ITERATES III

By

S.L. SINGH*
Department of Mathematics
L.M.S. Government College
Rishikesh, Dehra Dun 249201 India

It is well-known that the Picard sequence of iterates of a non-expansive mapping T from a closed & bounded subset C of a convex Banach space into itself need not converge to a fixed point of T. Therefore, for each point x in C, one considers a Picard sequence starting from x and generated by T_t where $T_t\, x = t\, Tx + (1-t)\, x$, $t \in (0, 1)$. The intent of this note is to present some convergence theorems for such sequences, which generalize several known results.

1.

In this section two results related with the convergence of the sequence of iterates of a *densifying* mapping to a fixed point are given. For a densifying mapping, one may see Singh [9] & [10]. It is also called "condensing" (*e.g.* see Ray [4]).

Motivated by a general type of mapping introduced by Hardy and Rogers [3], Singh [9] or [10] and Ray [4] independently proved the following result.

Theorem A.

Let C be a closed, bounded and convex subset of a strictly convex Banach space X. Let $T : C \to C$ be a densifying mapping satisfying

$$(1.1)\quad \| Tx - Ty \| \leq a \| x - y \| + b\,(\| x - Tx \| + \| y - Ty \|) + c\,(\| y - Tx \| + \| x - Ty \|),$$

for all x, y in C and for nonnegative numbers a, b, c with $a + 2b + 2c \leq 1$. Then, for $x_0 \in C$, the sequence of interates

* This work was partially supported by the UGC, New Delhi (Code No. 7574).

$\{ T_t^n x_0 \}$, where $T_t : C \longrightarrow C$ is defined by

$$T_t = t\, I + (1 - t)\, T, \; t \in (0, 1), \tag{1.2}$$

converges to a fixed point of T.

Several special cases of this theorem have been obtained. We mention a few. In case $b = c$, $a = 1$, $t = \frac{1}{2}$ we get a result due to Edelstein [2]. In case $c = 0$, we get a theorem due to Singh and Riggio [11]. In case $c = 0$ and T is completely continuous, we get a result due to Barbutti and Guerra [1]. In case $b = c = 0$, $a = 1$ and T is completely continuous, the result of Reinermann [5] is obtained.

Theorem A has been generalized by Singh [7]. Theorem 1 of this note presents another generalization. Let the mapping T on C be such that

$$\| Tx - Ty \| \leq k.\max \{ \| x - y \|, \\ \tfrac{1}{2} (\| x - Tx \| + \| y - Ty \|) + k' (\| x - Ty \| + \| y - Tx \|) \} \tag{1.3}$$

for all x, y in C and for $k, k' \geq 0$ with $k + 2k' \leq 1$.

It is clear that (1.1) $\Rightarrow$ (1.3), that is, mappings satisfying (1.1) also satisfy (1.3). The example in Theorem 1 (*xxv*) [6, p. 266] can be used to see that (1.3) does not imply (1.1). Hence Theorem A becomes indeed a special case of the following :

Theorem 1.

Let C be a closed, bounded and convex subset of a strictly convex Banach space X. Let $T : C \to C$ be a densifying mapping satisfying the condition (1.3). Then for each x_0 in C, the sequence of iterates $\{ T_t^n x_0 \}$, where $T_t : C \to C$ is a mapping defined by (1.2), converges to a fixed point of T.

Proof.

The result follows from a well known result (see, for instance, Singh [10, Theorem 3.3.1]), if we could show that

$\| T_t x - p \| < \| x - p \|$ for p in $F(T_t)$ and for all x in $C - F(T_t)$. Since T satisfies (1.3), we have

$$\| Tx-p \| = \| Tx-Tp \| \leq k.\ \max \{ \| x-p \|, \tfrac{1}{2} \| x-Tx \| \} + k' (\| x-p \| + \| p-Tx \|),$$

which implies

$$\| Tx-p \| \leq \| x-p \| . \tag{1.4}$$

A standard argument (see, for instance, [1, *p*. 30–31] or [9, *p*. 504]), using (1.4) and the strict convexity of X, shows that $\| T_t x-p \| < \| x-p \|$ for all x in $C-F(T_t)$ and p in $F(T_t)$.

Theorem 2.

With the hypotheses of Theorem 1 if, in place of (1.3), T satisfies the condition

$$\| Tx-Ty \| \leq k.\ \max \{ \| x-y \|, \tfrac{1}{2} (\| x-Ty \| + \| y-Tx \|) \} + k' (\| x-Tx \| + \| y-Ty \|), \tag{1.5}$$

for all x, y in C and for $k, k' \geqslant 0$ with $k+2k' \leq 1$, then $\left\{ T_t^n x_0 \right\}$ converges to a fixed point of T.

Proof

As in Theorem 1.

In her unpublished work, the condition (1.5) has recently been introduced by S. Ranganathan (B.H.U., Varanasi). It is easy to see that (1.1) $\Rightarrow$ (1.5). Thus Theorem 2 presents another generalization of Theorem A.

2.

In this section we consider the condition (1.3) and (1.5) with $k+2k' = 1$. With this special case while conditions on the space are relaxed considerably, we assume the convergence of $\{x_n\}$, where $x_{n+1} = (1-t)\, x_n + t\, Tx_n$.

Wong [12] proved the following :

Theorem B.

Let X be a convex subset of a normed linear space B. Let T be a self-mapping on X. Suppose that there exist a_i, $i = 1, \ldots, 5$ in [0, 1] such that $\sum_{i=1}^{5} a_i = 1$ and for all x, y in X,

$$\| Tx - Ty \| \leq a_1 \| x - Tx \| + a_2 \| y - Ty \| + a_3 \| x - Ty \| + a_4 \| y - Tx \| + a_5 \| x - y \| . \tag{2.1}$$

Let $x_0 \in X$, $t \in (0, 1)$ and $x_{n+1} = (1 - t)\, x_n + t \;\; Tx_n$ for each integer $n > 0$. Suppose that the sequence $\{ x_n \}$ converges to a point u in X. Then u is a fixed point of T.

This theorem has been generalized by Singh [8]. Two other generalizations are presented here. Note that (2.1) $\Rightarrow$ (1.3) and (2.1) $\Rightarrow$ (1.5) with $k + 2k' = 1$.

Theorem 3.

Let X be a convex subset of a normed linear space B. Let T be a self-mapping of X. Suppose that there exist k in [0, 1], k' in (0, 1) such that $k + 2k' = 1$ and for all x, y in X, (1.3) holds. Let $x_0 \in X$, $t \in (0, 1)$ and $x_{n+1} = (1 - t)\, x_n + t\; Tx_n$ for each integer $n \geq 0$. Suppose that the sequence $\{x_n\}$ converges to a point u in X. Then $Tu = u$.

Proof

Let $n \geq 0$. Then

(2.2) $$\| x_{n+1} - Tu \| = \| (1 - t)(x_n - Tu) + t\,(Tx_n - Tu) \| \leq (1 - t) \| x_n - Tu \| + t \| Tx_n - Tu \| .$$

By hypothesis,

(2.3) $$\| Tx_n - Tu \| \leq k . \max \{ \| x_n - u \| , \tfrac{1}{2} (\| x_n - Tx_n \| + \| u - Tu \|) \} + k' (\| x_n - Tu \| + \| u - Tx_n \|).$$

Since $\{x_m\}$ converges to u and $x_{n+1} - x_n = t\,(Tx_n - x_n)$, $\{Tx_m - x_m\}$ converges to 0. Using (2.3), (2.2) becomes by letting $n \to \infty$,

$$\| u - Tu \| \leq (1 - t) \| u - Tu \| + t\,(k \| u - Tu \| + k' \| u - Tu \|)$$

(2.4) $$= ((1 - t) + t\,(k + k'))\, \| u - Tu \| .$$

By assumption $0 < k + k' < 1$. So $0 < 1 - t + t(k + k') < 1$. Hence from (2.4), $Tu = u$.

Theorem 4.

With the hypotheses of Theorem 3, if in place of (1.3), T satisfies the condition (1.5) with $k, k' \in [0, 1]$ and $k + 2k' = 1$ then u is a fixed point of T.

Proof

Let $n \geq 0$. Then by (1.5),

$$\|Tx_n - Tu\| \leq k.\max\left\{\|x_n - u\|, \tfrac{1}{2}\left(\|x_n - Tu\| + \|u - Tx_n\|\right)\right\} + k'\left(\|x_n - Tx_n\| + \|u - Tu\|\right). \tag{2.5}$$

Recalling that (2.2) holds and the sequence $\{Tx_m - x_m\}$ and $\{u - Tx_m\}$ both converge to 0, we have, from (2.2) and (2.5), by letting $n \to \infty$,

$$\begin{aligned}\|u - Tu\| &\leq (1 - t)\|u - Tu\| + t(k/2 + k')\|u - Tu\| \\ &= (1 - t + t(k + 2k')/2)\|u - Tu\| \\ &= (1 - t/2)\|u - Tu\|.\end{aligned}$$

Since $0 < 1 - t/2 < 1$, $u = Tu$.

We thank the referee for his kind suggestions to improve the manuscript.

REFERENCES

1. U. Barbutti and S.G. Guerra, On an extension of a theorem of J.B. Diaz and F.T. Metcalf, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. 51 (1971), 29 – 31.

2. M. Edelstein, A remark on a theorem of M.A. Krasnoselskii, Amer. Math. Monthly 73 (1966), 509 – 510.

3. G.E. Hardy and T.D. Rogers A generalization of a fixed point theorem of Reich, Canad. Math. Bull. 16 (1973), 201-206.

4. B.K. Ray, Some results on fixed points, Indian J. Pure Appl. Math. 6 (1975), 410-415.

5. J. Reinermann, On a fixed point theorem of Banach type, Ann. Polon. Math. 23 (1970-71), 105-107.

6. B.E. Rhoades, A comparison of various definitions of contractive mappings, Trans. Amer. Math. Soc. 226 (1977), 257-290.

7. S.L. Singh, A note on the convergence of sequence of iterates, J. Natur. Sci. Math. 17 (1) (1977), 69-71.

8. S.L. Singh, A note on the convergence of sequence of iterates II, ibid. 17 (2) (1977), 15-17.

9. S.P. Singh, On the convergence of sequence of iterates, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. 57 (1974), 502-505.

10. S.P. Singh, Lecture Notes on Fixed Point Theorems in Metric and Banach Spaces, Matscience, Madras, (1974).

11. S.P. Singh and M.I. Riggio, On a theorem of Diaz and Metcalf, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Math. Natur. 53 (1972), 246-269.

12. C.S. Wong, Fixed points and characterization of certain maps, Pacific J. Math. 54 (1974), 305-312.

The Punjab University
Journal of Mathematics
Vol. XIV—XV (1981-82), pp. 129-134

A COMMON FIXED POINT THEOREM FOR OPERATORS ON BANACH SPACES

By

S L. SINGH

Department of Mathematics, L.M.S.
Government Postgraduate College,
Rishikesh, Dehra Dun, 249201, India.

A result generalizing fixed point theorems of Lernfeld, Lakshmikantham and Reddy, and Naik is presented.

1. Introduction

Let E be a Banach space and E_0 be the Banach space of all continuous functions from a finite closed interval I to E where

$$\| f \|_{E_0} = \sup_{t \in I} \| f(t) \|_E, \forall f \in E_0 .$$

Let S and T be operators from E_0 to E. An element $f \in E_0$ is said to be a fixed point [1] of T if

$$Tf = f(c) \text{ for some fixed } c \in I.$$

Consider the following conditions :

$$\| Tf - Tg \|_E \leq k \| f - g \|_{E_0} \quad \text{... (1)}$$

for all $f, g \in E_0$ where $k \in [0, 1)$.

$$\| Tf - Tg \|_E \leq p \| f - g \|_{E_0} + q (\| f(c) - Tf \|_E + \| g(c) - Tg \|_E) + r (\| f(c) - Tg \|_E + \| g(c) - Tf \|_E) \quad \text{... (2)}$$

for all $f, g \in E_0$ where $p, q\ r \geq 0$, $p + 2q + 2r < 1$.

The conditions (1) and (2) have been introduced by Bernfeld at al. [1] and Naik [2] respectively. We introduce the following :

$$\| Sf - Tg \|_E \le k . \max \{ \| f - g \|_{E_{E_0}} , \| f(c) - Sf \|_E , \| g(c) - Tg \|_E , [\| g(c) - Sf \|_E + \| f(c) - Tg \|_E] / 2 \} \quad \ldots (3)$$

for all $f, g \in E_0$ where $k \in [0, 1)$.

We remark that (1) $\Rightarrow$ (2) $\Rightarrow$ (3) if S = T in (3), *i.e.*, operators satisfying the contraction condition (1) will satisfy (2) and those satisfying (2) will also satisfy the special case (S=T) of (3). Result of this note is proved under the condition (3).

2. RESULT

Theorem

Suppose that S, T : $E_0 \to E$ satisfy (3). Then the following hold :

(*i*) Given $f_0 \in E_0$, every sequence of iterates $\{f_n\}$ satisfying $Sf_{2n} = f_{2n+1}(c)$, $Tf_{2n+1} = f_{2n+2}(c)$, $n = 0\ 1, 2, \ldots$, for a given $c \in I$ and $\| f_{n+1} - f_n \|_{E_0} = \| f_{n+1}(c) - f_n(c) \|_E$ converges to a common fixed point f^* of S and T.

(*ii*) Given f_0, $g_0 \in E_0$, let $\{f_n\}$ and $\{g_n\}$ be the sequences of iterates corresponding to f_0 and g_0 constructed as in (*i*). Then

$$\| f_n - g_n \|_{E_0} \le (1 - k)^{-1} (\| f_1 - f_0 \|_{E_0} + \| g_1 - g_0 \|_{E_0}) + \| f_0 - g_0 \|_{E_0} \quad \ldots (4)$$

If, in particular, $f_0 = g_0$ and $\{f_n\} \neq \{g_n\}$, then

$$\| f_n - g_n \|_{E_0} \le 2 (1 - k)^{-1} \| f_1 - f_0 \|_{E_0} .$$

(*iii*) Let $K_0 = \{f \in E_0 : \| f \|_{E_0} = \| f(c) \|_E\}$ and let $\{f_n\}$ and $\{g_n\}$ be as in (*ii*). If $f_n - g_n \in K_0$ for all n, then

$$\lim f_n = \lim g_n \qquad \dots (5)$$

Finally, if we define

$$K_{f^*} = \{f \in E_0 : \| f - f^* \|_{E_0} = \| f(c) - f^*(c) \|_E\}$$

where f^* is a common fixed point of S and T, then f^* is the only common fixed point of S and T in K_{f^*}.

Proof

Let $f_0 \in E_0$. Choose $\{f_n\}$ as in (*i*). By (3),

$$\begin{aligned} \| f_{2n+1} - f_{2n+2} \|_{E_0} &= \| f_{2n+1}(c) - f_{2n+2}(c) \|_E \\ &= \| Sf_{2n} - Tf_{2n+1} \|_E \\ &\leq k.\max\{ \| f_{2n} - f_{2n+1} \|_{E_0}, \\ &\quad \tfrac{1}{2} \| f_{2n}(c) - f_{2n+2}(c) \|_E \}. \end{aligned}$$

This gives $\| f_{2n+1} - f_{2n+2} \|_{E_0} \leq k \| f_{2n} - f_{2n+1} \|_{E_0}$.

Similarly $\| f_{2n+2} - f_{2n+3} \|_{E_0} \leq k \| f_{2n+1} - f_{2n+2} \|_{E_0}$.

Hence $\| f_n - f_{n+1} \|_{E_0} \leq k \| f_{n-1} - f_n \|_{E_0}$, $n = 1, 2, 3 \dots$

Consequently, $\{f_n\}$ is a Cauchy sequence and converges to some $f^* \in E_0$. By (3),

$$\begin{aligned} \| Sf^* - Tf_{2n+1} \|_E \leq k.\max\{ &\| f^* - f_{2n+1} \|_{E_0}, \\ &\| f^*(c) - Sf^* \|_E, \| f_{2n+1}(c) - Tf_{2n+1} \|_E, \\ &\tfrac{1}{2}[\| f_{2n+1}(c) - Sf^* \|_E + \| f^*(c) - Tf_{2n+1} \|_E] \}, \end{aligned}$$

that is

$$\| Sf^* - f_{2n+2}(c) \|_E \leq k. \max \{ \| f^* - f_{2n+1} \|_{E_0},$$
$$\| f^*(c) - Sf^* \|_E, \| f_{2n+1}(c) - f_{2n+2}(c) \|_E$$
$$\tfrac{1}{2} [\| f_{2n+1}(c) - Sf^* \|_E + \| f^*(c) - f_{2n+2}(c) \|_E]\}.$$

Making $n \to \infty$, we obtain

$$\| Sf^* - f^*(c) \|_E \leq k \| f^*(c) - Sf^* \|_E,$$

implying $Sf^* = f^*(c)$. Similarly $Tf^* = f^*(c)$. This proves (*i*).

Proof of (4) is the same as in [2]. If $g_0 = f_0$ then $g_0(c) = f_0(c)$ and $S g_0 = S f_0$. So $g_1(c) = f_1(c)$. Hence from (4),

$$\| f_n - g_n \|_{E_0} \leq (1-k)^{-1} (\| f_1 - f_0 \|_{E_0}$$
$$+ \| f_1(c) - f_0(c) \|_E)$$
$$= 2(1-k)^{-1} \| f_1 - f_0 \|_{E_0},$$

proving (*ii*).

To prove (*iii*) we note that $f_n - g_n \in K_0$ for $n = 0, 1, 2, \ldots \ldots$

$$\| f_{2n+1} - g_{2n+1} \|_{E_0} = \| f_{2n+1}(c) - g_{2n+1}(c) \|_E$$
$$\leq \| f_{2n+1}(c) - g_{2n+2}(c) \|_E$$
$$+ \| g_{2n+2}(c) - g_{2n+1}(c) \|_E$$
$$= \| Sf_{2n} - T g_{2n+1} \|_E$$
$$+ \| g_{2n+2} - g_{2n+1} \|_{E_0}$$
$$\leq k. \max \{ \| f_{2n} - g_{2n+1} \|_{E_0},$$
$$\| f_{2n}(c) - f_{2n+1}(c) \|_E,$$
$$\| g_{2n+1}(c) - g_{2n+2}(c) \|_E,$$

$$
\begin{aligned}
&\tfrac{1}{2}[\ \| g_{2n+1}(c) - f_{2n+1}(c) \|_E \\
&+ \| f_{2n}(c) - g_{2n+2}(c) \|_E \ ]\} \\
&+ \| g_{2n+2} - g_{2n+1} \|_{E_0} \\
&\leq k.\ \max\{\ \| f_{2n} - g_{2n+1} \|_{E_0}, \\
&\quad \| f_{2n} - f_{2n+1} \|_{E_0}, \\
&\quad \| g_{2n+1} - g_{2n+2} \|_{E_0}, \\
&\tfrac{1}{2}[\ \| g_{2n+1} - f_{2n+1} \|_{E_0} \\
&+ \| f_{2n} - f_{2n+1} \|_{E_0} \\
&+ \| f_{2n+1} - f_{2n+2} \|_{E_0} \\
&+ \| f_{2n+2} - g_{2n+2} \|_{E_0} \ ]\} \\
&+ \| g_{2n+2} - g_{2n+1} \|_{E_0}.
\end{aligned}
$$

Making $n \to \infty$ we obtain

$$\| f^* - g^* \|_{E_0} \leq k \| f^* - g^* \|_{E_0},$$

Since $f_n \to f^*$ and $g_n \to g^*$, where f^* and g^* are common fixed points of S and T. So $f^* = g^*$ and this proves (5).

If $f^*, g^* \in K_{f^*}$ are distinet common fixed points of S and T, then by (3),

$$
\begin{aligned}
\| f^* - g^* \|_{E_0} &= \| f^*(c) - g^*(c) \|_E \\
&= \| Sf^* - Tg^* \|_E \\
&\leq k \| f^* - g^* \|_{E_0},
\end{aligned}
$$

a contradiction. Hence S and T have a unique common fixed point in K_{f*}.

Bernfeld et al. [1] proved the above theorem under the contraction condition (1), and Naik [2] generalized their result by considering the condition (2). Thus, in view of the remark preceding the above theorem, the above theorem presents an interesting generalization of their results.

REFERENCES

1. S. Bernfeld, V. Lakshmikantham and Y. M. Reddy, Fixed point theorems of operators with PPF dependence in Banach spaces, J. Applicable Analysis 6 (1977), 271-280.

2. K. Viswanatha Naik, Fixed point theorems for operators on Banach spaces, Indian J. Pure Appl. Math. 12 (1981), 312-316.

The Punjab University
Journal of Mathematics
Vol. XIV—XV (1981-82), pp. 135-154.

THE INVERSE SPECTRAL PROBLEM FOR A SYSTEM OF THREE COUPLED FIRST ORDER EQUATIONS

By

P. J. CAUDREY

Department of Mathematics, UMIST,
P.O. Box 88, Sackville Street
Manchester M 60 IQD, U.K.

and

MUHAMMAD KALIM

Department of Mathematics
Government College, Lower Mall,
Lahore-Pakistan

Abstract

The spectral transform has proved to be a powerful tool for solving non-linear evolution equations. In this paper we consider a spectral problem. We define a set of spectral data and show how we can recover the potentials from this data. As an example the non-linear Klein-Gordon equation is solved by using the spectral transform.

§ 1. Introduction

In this paper the spectral problem

$$
\begin{aligned}
u_{1_x} - i\zeta u_1 &= q_{12}(x)\, u_2 + q_{13}(x)\, u_3 \\
u_{2_x} - i w^2 \zeta u_2 &= q_{23}(x)\, u_3 + q_{21}(x)\, u_1 \qquad (1) \\
u_{3_x} + i w \zeta u_3 &= q_{31}(x)\, u_1 + q_{32}(x)\, u_2
\end{aligned}
$$

where $w = e^{i\pi/3}$ is considered and solved by using the Spectral Transform. A set of spectral data which is sufficient for the reconstruction of the potentials $q_{12}(x)$, $q_{13}(x)$, $q_{23}(x)$, $q_{21}(x)$, $q_{31}(x)$, $q_{32}(x)$ is found, and the problem of this reconstruction, the inverse problem is solved.

However, until recently the only spectral problems for which the inverse problem had been solved were second order ones by Ablowitz [1] et al. and their matrix generalisations by Wadati [2] and Calogero and Degasperis.

As an example the non-linear Klein-Gordon equation

$$\sigma_{xt} = e^{26} - e^{-6}$$

is solved by using spectral transform.

§ 2. The Direct Spectral Problem

We can define Jost functions which are solutions of (1) by the boundary conditions

$$e^{-i\zeta x}\, \underline{\phi}_1(x, \zeta) \to \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{as } x \to -\infty \tag{2a}$$

$$e^{-i\omega^2\zeta x}\, \underline{\phi}_2(x, \zeta) \to \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{as } x \to -\infty \tag{2b}$$

$$e^{i\omega\zeta x}\, \underline{\phi}_3(x, \zeta) \to \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{as } x \to -\infty \tag{2c}$$

and

$$e^{-i\zeta x}\, \underline{\psi}_1(x, \zeta) \to \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{as } x \to +\infty \tag{3a}$$

$$e^{-i\alpha^2\zeta x}\,\underline{\psi}_2(x,\zeta) \to \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{as } x \to +\infty \tag{3b}$$

$$e^{i\alpha^2\zeta x}\,\underline{\psi}_3(x,\zeta) \to \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{as } x \to +\infty \tag{3c}$$

Together with the boundary condition (2*a*) equation (1) is equivalent to

$$e^{-i\zeta x}\phi_{11}(x,\zeta) = 1 +$$

$$\int_{-\infty}^{x} \left[q_{12}(y)\, e^{-i\zeta y}\phi_{12}(y,\zeta) + q_{13}(y)\, e^{-i\zeta y}\phi_{13}(y,\zeta) \right] dx$$

$$e^{-i\zeta x}\phi_{12}(x,\zeta) = \int_{-\infty}^{x} e^{-\sqrt{3}\, w\zeta (x-y)}$$

$$\left\{ \{q_{23}(y)\, e^{-i\zeta y}\phi_{13}(y,\zeta) + q_{21}(y)\, e^{-i\zeta y}\phi_{11}(y,\zeta)\} \right\} dy$$

$$e^{-i\zeta x}\phi_{13}(x,\zeta) = \int_{-\infty}^{x} e^{-\sqrt{3}\,\omega^2\zeta (x-y)}$$

$$\left\{ \{q_{31}(y)\, e^{-i\zeta y}\phi_{11}(y,\zeta) + q_{32}(y)\, e^{-i\zeta y}\phi_{12}(y,\zeta)\} \right\} dy \tag{4}$$

Expanding $e^{-i\zeta x}\phi_1(x,\zeta)$ as a Neumann series

$$e^{-i\zeta x}\underline{\phi}_1(x,\zeta) = \sum_{n=o}^{\infty} \mathbf{h}_1^{n}(x,\zeta)$$

$$\mathbf{h}_1^{(o)}(x,\zeta) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$h_{11}^{(n+1)}(x,\zeta) = \int_{-\infty}^{x} \{ q_{12}(y)\, h_{12}^{(n)}(y,\zeta) + q_{13}(y)\, h_{13}^{(n)}(y,\zeta) \}\, dy$$

$$h_{12}^{(n+1)}(x,\zeta) = \int_{-\infty}^{x} e^{-\sqrt{3}\, w \zeta (x-y)} \{ q_{23}(y)\, h_{13}^{(n)}(y,\zeta) + q_{21}(y)\, h_{11}^{(n)}(y,\zeta) \}\, dy$$

$$h_{13}^{(n+1)}(x,\zeta) = \int_{-\infty}^{x} e^{-\sqrt{3}\, \omega^2 \zeta (x-y)} \{ q_{31}(y)\, h_{11}^{(n)}(y,\zeta) + q_{32}(y)\, h_{12}^{(n)}(y,\zeta) \}\, dy \tag{5}$$

If we define

$$Q(x) = \int_{-\infty}^{x} \max \{ q_{ij}(y) \}\, dy, \qquad Q(\infty) < \infty$$

then we can show that

$$\left.\begin{aligned} | h_{11}^{(n)}(x,\zeta) | &\le \frac{1}{3} \left\{ 2^n + 2(-1)^n \right\} \frac{1}{n!} Q^n(x) \\ | h_{\substack{12\\13}}^{(n)}(x,\zeta) | &\le \frac{1}{3} \left\{ 2^n - (-1)^n \right\} \frac{1}{n!} Q^n(x) \end{aligned}\right]$$

$$-\frac{5\pi}{6} \le \arg \zeta \le -\frac{\pi}{6}$$

Thus the Neumann series converges for $-\frac{5\pi}{6} \le \arg \zeta \le -\frac{5}{6}$ by comparison with the series

$$\sum_{n=0}^{\infty} \frac{1}{3} \left\{ 2^n + 2(-1)^n \right\} \frac{1}{n!} Q^n(x) = \frac{1}{3} \exp\{2Q(x)\} + \frac{2}{3} \exp\{-Q(x)\}$$

and

$$\sum_{n=0}^{\infty} \frac{1}{3}\left(2^n - (-1)^n\right)\frac{1}{n!}\,Q^n(x) = \frac{1}{3}\exp\{2\,Q(x)\}$$

$$-\frac{1}{3}\exp\{-Q(x)\}$$

It follows that

$$\left| e^{-i\zeta x}\,\phi_{11}(x,\zeta)\right| \leq \frac{1}{3}\exp\{2\,Q(x)\} + \frac{2}{3}\exp\{-Q(x)\}$$

and

$$\left| e^{-i\zeta x}\,\phi_{\substack{12\\13}}(x,\zeta)\right| \leq \frac{1}{3}\exp\{2\,Q(x)\} - \frac{1}{3}\exp\{-Q(x)\}$$

Thus $\underline{\phi}_1(x,\zeta)$ is uniquely defined for

$$-\frac{5\pi}{6} \leq \arg\zeta \leq -\frac{\pi}{6}$$

and is analytic in the interior of this region. Also by using the Riemann-Lesbegue lemma it can be seen that

$$e^{-i\zeta x}\,\underline{\phi}_1(x,\zeta) \to \begin{bmatrix}1\\0\\0\end{bmatrix}$$

as $\zeta \to \infty$ in the region.

Similar results apply for $\underline{\phi}_2(x,\zeta)$ and $\underline{\phi}_3(x,\zeta)$ which are defined

for

$$\frac{\pi}{2} \leq \arg\zeta \leq \frac{7\pi}{6} \text{ and } -\frac{\pi}{6} \leq \arg\zeta \leq \frac{\pi}{2} \text{ respectively.}$$

$$e^{-iw^2\zeta x}\,\underline{\phi}_2(x,\zeta) \to \begin{bmatrix}0\\1\\0\end{bmatrix} \text{ and } e^{-iw\zeta x}\,\underline{\phi}_3(x,\zeta) \to \begin{bmatrix}0\\0\\1\end{bmatrix}$$

as $x \to -\infty$

Also we have $\psi_1(x,\zeta)$, $\psi_2(x,\zeta)$ and $\psi_3(x,\zeta)$ as defined by (3) in the regions

$\frac{\pi}{6} \le \arg\zeta \le \frac{5\pi}{6}$, $-\frac{\pi}{2} \le \arg\zeta \le \frac{\pi}{6}$ and $\frac{5\pi}{6} \le \arg\zeta \le \frac{3\pi}{2}$ respectively we need more Jost functions.

These can be defined as follows. First we find the Jost functions

$\underline{\phi}_2^+(x,\zeta)$, $\underline{\phi}_2^+(x,\zeta)$, $\underline{\phi}_3^+(x,\zeta)$, $\underline{\psi}_1^+(x,\zeta)$ $\underline{\psi}_2^+(x,\zeta)$ and

$\underline{\psi}_3(x,\zeta)$ of the conjugate spectral problem.

$$\overset{+}{V}_{1x} + i\zeta \overset{+}{V}_1 = -q_{21}(x)\,\overset{+}{V}_2 - q_{31}(x)\,\overset{+}{V}_3$$

$$\overset{+}{V}_{2x} + i w^2 \zeta\,\overset{+}{V}_2 = -q_{32}(x)\,\overset{+}{V}_3 - q_{12}(x)\,\overset{+}{V}_1$$

$$\overset{+}{V}_{3x} - i w \zeta\,\overset{+}{V}_3 = -q_{13}(x)\,\overset{+}{V}_1 - q_{23}(x)\,\overset{+}{V}_2 \qquad (6)$$

Using the same methods as before we get

$$e^{i\zeta x}\,\underline{\phi}_1^+(x,\zeta) \rightarrow \begin{bmatrix}1\\0\\0\end{bmatrix} \text{ as } x\rightarrow -\infty, \text{ defined for } \frac{\pi}{6} \le \arg\zeta \le \frac{5\pi}{6} \qquad (7\text{ a})$$

$$e^{i w^2\zeta x}\,\underline{\phi}_2^+(x,\zeta) \rightarrow \begin{bmatrix}0\\1\\0\end{bmatrix} \text{ as } x\rightarrow -\infty, \text{ defined for } -\frac{\pi}{2} \le \arg\zeta \le \frac{\pi}{6} \qquad (7\text{b})$$

$$e^{-i w\zeta x}\,\underline{\phi}_3^+(x,\zeta) \quad \begin{bmatrix}0\\0\\1\end{bmatrix} \text{ as } x\rightarrow -\infty, \text{ defined for } \frac{5\pi}{6} \le \arg\zeta \le \frac{3\pi}{2} \qquad (7\text{c})$$

$$e^{i\zeta x}\,\underline{\psi}_1^{+}(x,\zeta) \to \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \text{ as } x \to +\infty, \text{ defined for}$$

$$-\frac{5\pi}{6} \le \arg \zeta \le -\frac{\pi}{6} \quad (8a)$$

$$e^{i w^2 \zeta x}\,\underline{\psi}_2^{+}(x,\zeta) \to \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ as } x \to +\infty, \text{ defined for}$$

$$\frac{\pi}{6} \le \arg \zeta \le \frac{7\pi}{6} \quad (8b)$$

$$e^{-i w \zeta x}\,\psi_3^{+}(x,\zeta) \to \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ as } x \to +\infty, \text{ defined for}$$

$$-\frac{\pi}{6} \le \arg \zeta \le \frac{\pi}{2} \quad (8c)$$

We can now define

$$\underline{\theta}_1(x,\zeta) = \underline{\phi}_2^{+}(x,\zeta) \times \underline{\psi}_3^{+}(x,\zeta), \text{ defined for } -\frac{\pi}{6} \le \arg\zeta \le \frac{\pi}{6} \quad (9a)$$

$$\underline{\theta}_2(x,\zeta) = \underline{\phi}_3^{+}(x,\zeta) \times \underline{\psi}_1^{+}(x,\zeta), \text{ defined for } -\frac{5\pi}{6} \le \arg\zeta \le -\frac{\pi}{2} \quad (9b)$$

$$\underline{\theta}_3(x,\zeta) = \underline{\phi}_1^{+}(x,\zeta) \times \underline{\psi}_2^{+}(x,\zeta), \text{ defined for } \frac{\pi}{2} \le \arg\zeta \le \frac{5\pi}{6} \quad (9c)$$

$$\underline{\rho}_1(x,\zeta) = \underline{\psi}_2^{+}(x,\zeta) \times \underline{\phi}_3(x,\zeta), \text{ defined for } \frac{5\pi}{6} \le \arg\zeta \le \frac{7\pi}{6} \quad (10a)$$

$$\underline{\rho}_2(x,\zeta) = \underline{\psi}_3^{+}(x,\zeta) \times \underline{\phi}_1^{+}(x,\zeta), \text{ defined for } \frac{\pi}{6} \le \arg\zeta \le \frac{\pi}{2} \quad (10b)$$

$$\rho_3(x, \zeta) = \psi_1^{+}(x, \zeta) \times \phi_2^{+}(x, \zeta), \text{ defined for } -\frac{\pi}{2} \leq \arg \zeta \leq -\frac{\pi}{6} \quad \text{(10c)}$$

where $\times$ is the standard vector product. If

$$\int_{-\infty}^{\infty} \max \{ | q_{ij}(y) | \} e^{K|y|} dy < \infty \quad (11)$$

for all K then all functions are analytic throughout the complex ζ – plane. In this case we can define a matrix $\{a_{ij}(\zeta)\}$ by

$$\psi_i(x, \zeta) = \sum_{j=1}^{3} a_{ij}(\zeta) \phi_j(x, \zeta) \quad (12)$$

and its inverse $\{\bar{a}_{ij}(\zeta)\}$

$$\phi_i(x, \zeta) = \sum_{j=1}^{3} \bar{a}_{ij}(\zeta) \psi_i(x, \zeta) \quad (13)$$

Also we can easily show that

$$\det \{a_{ij}(\zeta)\} = \det \{\bar{a}_{ij}(\zeta)\} = 1,$$

$$\psi_i^{+}(x, \zeta) = \sum_{j=1}^{3} \bar{a}_{ji}(\zeta) \psi_j^{+}(x, \zeta) \quad (14)$$

and

$$\phi_i^{+}(x, \zeta) = \sum_{j=1}^{3} a_{ij}(\zeta) \psi_j^{+}(x, \zeta) \quad (15)$$

We now define $\mathbf{F}_1(x, \zeta)$, $\mathbf{F}_2(x, \zeta)$ and $\mathbf{F}_3(x, \zeta)$ which are the key functions in the inverse spectral problem

$$
e^{i\zeta x}\,\mathbf{F}_1(x,\zeta)=\begin{cases}
\underline{\phi}_1(x,\zeta) & \text{for } -\dfrac{5\pi}{6}<\arg\zeta<-\dfrac{\pi}{6}\\[2ex]
\dfrac{\underline{\theta}_1(x,\zeta)}{\bar{a}_{33}(\zeta)} & \text{for } -\dfrac{\pi}{6}<\arg\zeta<\dfrac{\pi}{6}\\[2ex]
\dfrac{\underline{\psi}_1(x,\zeta)}{\underline{a}_{11}(\zeta)} & \text{for } \dfrac{\pi}{6}<\arg\zeta<\dfrac{5\pi}{6}\\[2ex]
\dfrac{\underline{\rho}_1(x,\zeta)}{\bar{a}_{22}(\zeta)} & \text{for } \dfrac{5\pi}{6}<\arg\zeta<\dfrac{7\pi}{6}
\end{cases}
$$

Across the boundaries between the various regions we have the differences.

$$
\frac{\underline{\theta}_1(x,\zeta)}{\bar{a}_{33}(\zeta)}-\underline{\phi}_1(x,\zeta)=-\frac{\bar{a}_{13}(\zeta)}{\bar{a}_{33}(\zeta)}\,\underline{\phi}_3(x,\zeta)\ \text{ on } \arg\zeta=-\frac{\pi}{6}
$$

$$
\frac{\underline{\rho}_1(x,\zeta)}{\bar{a}_{22}(\zeta)}-\underline{\phi}_1(x,\zeta)=-\frac{\bar{a}_{12}(\zeta)}{\bar{a}_{22}(\zeta)}\,\underline{\phi}_2(x,\zeta)\ \text{ on } \arg\zeta=-\frac{5\pi}{6}
$$

$$
\frac{\underline{\psi}_1(x,\zeta)}{\underline{a}_{11}(\zeta)}-\frac{\underline{\theta}_1(x,\zeta)}{\bar{a}_{33}(\zeta)}=\frac{a_{12}(\zeta)}{a_{11}(\zeta)\,\bar{a}_{33}(\zeta)}\,\underline{\rho}_2(x,\zeta)\ \text{ on } \arg\zeta=\frac{\pi}{6}
$$

$$
\frac{\underline{\psi}_1(x,\zeta)}{a_{11}(\zeta)}-\frac{\underline{\rho}_1(x,\zeta)}{\bar{a}_{22}(\zeta)}=\frac{a_{13}(\zeta)}{a_{11}(\zeta)\,\bar{a}_{22}(\zeta)}\,\underline{\theta}_3(x,\zeta)\ \text{ on } \arg\zeta=\frac{5\pi}{6}
$$

also

$$e^{i w^2 \zeta x} \mathbf{F}_2(x, \zeta) = \begin{cases} \phi_2(x, \zeta) & \text{for } \frac{\pi}{2} < \arg \zeta < \frac{7\pi}{6} \\ \dfrac{\theta_2(x, \zeta)}{\bar{a}_{11}(\zeta)} & \text{for } -\frac{5\pi}{6} < \arg \zeta < -\frac{\pi}{2} \\ \dfrac{\psi_2(x, \zeta)}{a_{22}(\zeta)} & \text{for } -\frac{\pi}{2} < \arg \zeta < \frac{\pi}{6} \\ \dfrac{\rho_2(x, \zeta)}{\bar{a}_{33}(\zeta)} & \text{for } \frac{\pi}{6} < \arg \zeta < \frac{\pi}{2} \end{cases} \tag{17}$$

and

$$e^{-i w \zeta x} \mathbf{F}_3(x, \zeta) = \begin{cases} \phi_3(x, \zeta) & \text{for } -\frac{\pi}{6} < \arg \zeta < \frac{\pi}{2} \\ \dfrac{\theta_3(x, \zeta)}{\bar{a}_{22}(\zeta)} & \text{for } \frac{\pi}{2} < \arg \zeta < \frac{5\pi}{6} \\ \dfrac{\psi_3(x, \zeta)}{a_{33}(\zeta)} & \text{for } \frac{5\pi}{6} < \arg \zeta < \frac{3\pi}{2} \\ \dfrac{\rho_3(x, \zeta)}{\bar{a}_{11}(\zeta)} & \text{for } -\frac{\pi}{2} < \arg \zeta < -\frac{\pi}{6} \end{cases} \tag{18}$$

We also define (for real ξ)

$$f_1(x, -\xi) = \mathbf{F}_1(x, -i\,w^2(\xi - 10)) =$$

$$\begin{cases} \dfrac{\underline{\theta}_1(x, -i\,w^2\xi)}{\bar{a}_{33}(-i\,w^2\xi)}\, e^{-w^2\xi x} & \text{for } \xi > 0 \\ \underline{\phi}_1(x, -i\,w^2\xi)\, e^{-w^2\xi x} & \text{for } \xi < 0 \end{cases} \quad (19a)$$

$$f_2(x, -\xi) = \mathbf{F}_1(x, -i\,w(\xi - 10)) =$$

$$\begin{cases} \underline{\phi}_1(x, -i\,w\xi)\, e^{-\omega\xi x} & \text{for } \xi > 0 \\ \dfrac{\underline{\rho}_1(x, -i\,\omega\xi)}{\bar{a}_{22}(-i\,\omega\xi)}\, e^{-\omega\xi x} & \text{for } \xi < 0 \end{cases} \quad (19b)$$

Similarly

$$f_3(x, -\xi) = \mathbf{F}_2(x, -i\,\omega^2(\xi - 10)) \quad (20a)$$

$$f_4(x, -\xi) = \mathbf{F}_2(x, -i\,\omega(\xi - 10)) \quad (20b)$$

$$f_5(x, -\xi) = \mathbf{F}_3(x, -t\,\omega^2(\xi - 10)) \quad (21a)$$

and

$$f_6(x, -\xi) = \mathbf{F}_3(x, -i\,\omega(\xi - 10)) \quad (21b)$$

$\mathbf{F}_1(x, \zeta)$ has the following properties

(*i*) $\mathbf{F}_1(x, \zeta) \rightarrow \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ as $|\zeta| \rightarrow \infty$ (22)

(*ii*) on the boundary $\zeta = -i\,\omega\,\xi$ (ξ real)

$$\mathbf{F}_1 (x, - i \omega (\xi + 10)) = \mathbf{F}_1 (x, - i \omega (\xi - 10))$$

$$= R_{13} (- \xi) f_5 (x, - \xi) c^{- i \sqrt{3} \omega \xi x} \tag{23a}$$

where

$$R_{13} (\xi) = \begin{cases} \dfrac{- \bar{a}_{13} (i \omega^2 \xi)}{\bar{a}_{33} (i \omega^2 \xi)} & \text{for } \xi < 0 \\ \dfrac{a_{13} (i \omega^2 \xi)}{a_{11} (i \omega^2 \xi)} & \text{for } \xi > 0 \end{cases} \tag{23b}$$

Similarly on $\zeta = - i \omega^2 \xi$ (ξ real)

$$\mathbf{F}_1 (x, - i \omega^2 (\xi + io)) - \mathbf{F}_1 (x, - i \omega^2 (\xi - io))$$

$$= P_{12} (- \xi) f_4 (x, - \xi) e^{i \sqrt{3} \omega^2 \xi x} \tag{24a}$$

where

$$P_{12} (\xi) = \begin{cases} \dfrac{\bar{a}_{12} (i \omega \xi)}{\bar{a}_{22} (i \omega \xi)} & \text{for } \xi > 0 \\ - \dfrac{a_{12} (i \omega \xi)}{a_{11} (i \omega \xi)} & \text{for } \xi < 0 \end{cases} \tag{24b}$$

(*iii*) $\mathbf{F}_1 (x, \zeta)$ has poles (assumed simple) at the zeros

of $\bar{a}_{33} (\zeta)$, $a_{11} (\zeta)$ and $\bar{a}_{22} (\zeta)$ in the appropriate regions of the complex ζ - plane.

For $- \frac{\pi}{6} < \arg \zeta < \frac{\pi}{6}$ the poles occur

Where $\bar{a}_{33} (\zeta) = 0$ in which case if (11) is satisfied

$$\underline{\theta}_1 (x, \zeta) = - \bar{a}_{13} (\zeta) \underline{\phi}_3 (x, \zeta)$$

$$\text{Residue} = \frac{\underline{\theta}_1 (x, \zeta)}{\frac{d}{d\zeta} \bar{a}_{33} (\zeta)} e^{-i \zeta x}$$

$$= - \frac{\bar{a}_{13} (\zeta)}{\frac{d}{d\zeta} \bar{a}_{33} (\zeta)} \underline{\phi}_3 (x, \zeta)\, e^{- i \zeta x}$$

$$= \gamma\, \mathbf{F}_3 (x, \zeta)\, e^{- \sqrt{3}\, \omega^2 \zeta x} \qquad \text{(25a)}$$

this is a residue of second type. If (11) is not satisfied we proceed as follows.

$$\underline{\theta}_1 (x, \zeta) \times \underline{\phi}_3 (x, \zeta) = - \bar{a}_{33} (\zeta)\, \overset{+}{\underline{\phi}}_2 (x, \zeta) \qquad \text{(25b)}$$

Thus when $\bar{a}_{33} (\zeta) = 0$

$$\underline{\theta}_1 (x, \zeta) = \beta\, \underline{\phi}_3 (x, \zeta)$$

for some scalar β and (25*a*) follows without difficulty. Similarly for $\frac{5\pi}{6} < \arg \zeta < \frac{7\pi}{6}$ the poles occur where $\bar{a}_{22} (\zeta) = 0$. In this case

$$\text{Residue} = \beta\, \mathbf{F}_2 (x, \zeta)\, e^{- \sqrt{3}\, \omega \zeta x} \qquad \text{(25c)}$$

we shall call this is a residue of first type.

In this region $\frac{\pi}{6} < \arg \zeta < \frac{\pi}{2}$ the poles occur where $a_{11} (\zeta) = 0$ in which case if (11) holds

$$\underline{\rho}_2 (x, \zeta) = - a_{21} (\zeta)\, \underline{\psi}_1 (x, \zeta)$$

$$\underline{\psi}_1 (x, \zeta) = - \frac{\underline{\rho}_2 (x, \zeta)}{a_{21} (\zeta)}$$

$$\text{Residue} = \frac{\psi_1(x,\zeta)}{\frac{d}{d\zeta} a_{11}(\zeta)} e^{-i\zeta x}$$

$$= -\frac{\rho_2(x,\zeta)}{a_{21}(\zeta)\frac{d}{d\zeta} a_{11}(\zeta)} e^{-i\zeta x} \quad (\text{provided } a_{21}(\zeta) \neq 0)$$

$$= -\frac{\bar{a}_{33}(\zeta)}{a_{21}(\zeta)\frac{d}{d\zeta} a_{11}(\zeta)} \phi_2(x,\zeta) e^{-\sqrt{3}\,\omega\zeta x}$$

$$= \beta F_2(x,\zeta)\, e^{-\sqrt{3}\,w\zeta x}$$

which is a residue of first type. Thus provided $a_{21}(\zeta) \neq 0$, we have a residue of first type.

If $a_{21}(\zeta) = 0$ it follows that $\rho_2(x,\zeta) = 0$ then (using de l' Hospital rule).

$$F_2(x,\zeta) = \frac{\frac{\partial}{\partial\zeta}\rho_2(x,\zeta)}{\frac{d}{d\zeta}\bar{a}_{33}(\zeta)} e^{-i w^2 \zeta x}$$

It can easily be shown that

$$\psi_1(x,\zeta) \times \rho_2(x,\zeta) \cdot \phi_3(x,\zeta) = a_{11}(\zeta)\,\bar{a}_{33}(\zeta) \tag{27}$$

differentiating the equation (27) with respect to ζ and putting

$$a_{11}(\zeta) = \bar{a}_{33}(\zeta) = 0 \text{ and } \rho_2(x,\zeta) = 0$$

gives the result that $\psi_1(x,\zeta)$ is a linear combination of $\frac{d}{d\zeta}\rho_2(x,\zeta)$ and $\phi_3(x,\zeta)$. Therefore

$$\text{Residue} = \frac{1}{\frac{d}{d\zeta}(a_{11}(\zeta))} \psi_1(x,\zeta)\, e^{-i\zeta x}$$

$$= \beta \underline{\phi}_2 (x, \zeta)\, e^{-\sqrt{3}\, w \zeta x} + \gamma \underline{\phi}_3 (x, \zeta)\, e^{-3 w^2 \zeta x} \quad (28)$$

which is a linear combination of both types of residues.

Similar results apply for

$$\frac{\pi}{2} < \arg \zeta < \frac{5\pi}{6}$$

The functions

$$\underline{\phi}_2 (x, \zeta) \text{ and } \underline{\phi}_3 (x, \zeta)$$

behave in an analogous manner.

Spectral Data.

The spectral data required for solving the inverse problem are

$$S = \left\{ \begin{array}{l} R_{13} (\xi), P_{12} (\xi), R_{21} (\xi), P_{23} (\xi), R_{32} (\xi), P_{31} (\xi) \ (-\infty < \xi < \infty) \\ \zeta_K, \alpha_K, \beta_K, \gamma_K \ (K = 1, 2, \ldots\ldots\ldots n) \end{array} \right\}$$

where the ζ_K $(K = 1, 2, \ldots n)$ are the location of the poles of $\underline{\phi}_1 (x, \zeta)$, $\underline{\phi}_2 (x, \zeta)$ and $\underline{\phi}_3 (x, \zeta)$ and α_K, β_K γ_K are the co-efficients in the residues and $R_{13} (\xi)$, $P_{12} (\xi)$, $R_{21} (\xi)$, $P_{23} (\xi)$, $R_{32} (\xi)$, $P_{31} (\xi)$ $(-\infty\ \xi < \infty)$ from the continum part of the spectral data.

§ 3. The Inverse Problem

From the three properties of $F_1 (x, \zeta)$ we have

$$F_1 (x, \zeta_o) = 1 - \frac{1}{2\pi i} \int_{-\infty}^{\infty} \left\{ \frac{R_{13} (\xi)\, f_5 (x, \xi)}{\xi - i w^2 \zeta_o} e^{i \sqrt{3}\, w \xi x} + \frac{P_{12} (\xi)\, f_4 (x, \xi)}{\xi - i w \zeta_o} e^{-i \sqrt{3}\, w^2 \xi x} \right\} d\xi$$

$$- \sum_{K=1}^{n} \left[\frac{\gamma_K\, F_3 (x, \zeta_K)}{\zeta_K - \zeta_o} e^{-\sqrt{3}\, w^2 \zeta_K x} + \frac{\beta_K\, F_2 (x, \zeta_K)}{\zeta_K - \zeta_o} e^{-\sqrt{3}\, \omega \zeta_K x} \right] \ldots (29)$$

Putting $\zeta_o = - i\, w\, (\xi_o - io)$ in (29) gives

$$f_4 (x, \xi_o) = 1 - \frac{1}{2 \pi i} \int_{-\infty}^{\infty} \left(\frac{R_{13} (\xi)\, f_5 (x, \xi)}{\xi + \xi_o - io} e^{i \sqrt{3}\, w\, \xi\, x} + \frac{P_{12} (\xi)\, f_4 (x, \xi)}{\xi - w^2 \xi_o} e^{- i \sqrt{3}\, w^2 \xi x} \right) d \xi$$

$$- \sum_{K=1}^{n} \left[\frac{\gamma_K\, F_3 (x, \zeta_K)}{\zeta_K + i\, w\, \zeta_o} e^{- \sqrt{3}\, w^2 \zeta_K\, x} + \frac{\beta_K\, F_2 (x, \zeta_K)}{\zeta_K + i\, w\, \xi_o} e^{- \sqrt{3}\, w\, \zeta_K\, x} \right] \quad (30)$$

Putting $\zeta_o = i\, w^2 (\xi_o - io)$ in (29) gives

$$f_5 (x, \xi_o) = 1 - \frac{1}{2 \pi i} \int_{-\infty}^{\infty} \left\{ \frac{R_{13} (\xi)\, f_5 (x, \xi)}{\xi + w\, \xi_o} e^{i \sqrt{3}\, w\, \xi\, x} + \frac{P_{12} (\xi)\, f_4 (x, \xi)}{\xi + \xi_o - io} e^{- i \sqrt{3}\, w^2 \xi\, x} \right\} d\xi$$

$$- \sum_{K=1}^{n} \left[\frac{\gamma_K\, F_3 (x, \zeta_K)}{\zeta_K + i\, w^2 \xi_o} e^{- \sqrt{3}\, w^2 \zeta_K\, x} + \frac{\beta_K\, F_2 (x, \zeta_K)}{\zeta_K + i\, w^2 \xi_o} e^{- \sqrt{3}\, w\, \zeta_K\, x} \right] \quad (31)$$

Similarly we can find $f_1 (x, \xi_o)$, $f_2 (x, \xi_o)$, $f_3 (x, \xi_o)$ and $f_6 (x, \xi_o)$.

The functions $f_1 (x, \xi_o)$, $f_2 (x, \xi_o)$, $f_3 (x, \xi_o)$, $f_4 (x, \xi_o)$, $f_5 (x, \xi_o)$ and $f_6 (x, \xi_o)$, together with those found by putting $\zeta = - w\zeta_j, + \omega^2 \zeta_j$ $(j = 1, 2, \ldots n)$ form a set of singular. Fredholm/matrix equations which are linear in the unknowns

$$f_1 (x, \xi),\ f_2 (x, \xi),\ f_3 (x, \xi),\ f_4 (x, \xi),\ f_5 (x, \xi),\ f_6 (x, \xi),$$

$F_j (x, - w\, \zeta_K)$ and $F_j (x, w^2 \zeta_K)$, $\begin{Bmatrix} K = 1, 2, \ldots n \\ j = 1, 2, \ldots n \end{Bmatrix}$.

Thus these quantities can be found if the scattering data are given. Since each of them, when multiplied by an appropriate exponential of x, satisfies (1) the reconstruction of $q_{12}(x)$, $q_{13}(x)$, $q_{23}(x)$, $q_{21}(x)$, $q_{31}(x)$ and $q_{32}(x)$ are straightforward.

§ 4. **A Nonlinear Klein-Gordon Equation.**

One example of an evolution equation which is solvable by this spectral transform is the non-linear Klein-Gordon equation Fordy and Gibbons [3]

$$\sigma_{xt} = e^{26} - e^{-6} \tag{32}$$

we can write (1) as

$$U_x - i\zeta A \underset{-}{U} = Q(x) \underset{-}{U} \tag{33}$$

where **U** is a three-dimensional column vector and A and Q (x) are the 3×3 matrices.

$$A = \begin{bmatrix} \omega_1 & 0 & 0 \\ 0 & \omega_2 & \\ 0 & 0 & w_3 \end{bmatrix} \tag{34}$$

where $w_1 = 1$, $w_2 = w^2$, $\omega_3 = -w$

and

$$Q(x) = \begin{bmatrix} 0 & q_{12}(x) & q_{13}(x) \\ q_{21}(x) & 0 & q_{23}(x) \\ q_{31}(x) & q_{32}(x) & 0 \end{bmatrix} \tag{35}$$

If the matrix of potential Q (x, t) is a function of time t as well as x then the solution **U** (x, t, ζ) of (33) must also be time dependent.

If this dependence is given by

$$\mathbf{U}_t = P(x, t, \zeta)\, \mathbf{U} \tag{36}$$

cross-differentiation with (33) shows that

$$O_t = P_x + \{P, (i \zeta n + O)\} \tag{37}$$

Putting

$$O = \frac{i}{\sqrt{3}} 6_x \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix} \tag{38}$$

and

$$P = \frac{1}{3 i \zeta} (e^{26} + 2e^{-6}) \begin{bmatrix} \omega_1^{-1} & 0 & 0 \\ 0 & \omega_2^{-1} & 0 \\ 0 & 0 & \omega_1^{-1} \end{bmatrix} \tag{39}$$

makes (37) equivalent to (32), since (33) and (36) are satisfied by

$$\exp \{i \omega_j \zeta x + (i \omega_j \zeta)^{-1} t\} \mathbf{F}_j (x, t, \zeta) \qquad (j = 1, 2, 3) \tag{40}$$

Since (40) is a solution of (36) we can write

$$\left(- \frac{\partial}{\partial t} - P\right) \exp \{i \omega_j \zeta x + (i \omega_j \zeta)^{-1} t\} \mathbf{F}_j (x, t, \zeta) = 0 \tag{41}$$

we get the evolution of the spectral data

$$P_{12} (t, \zeta) = P_{12} (0, \xi) \exp (-\omega^2 \xi^{-1} t) \tag{42}$$

Similarly we get $P_{13} (t, \xi)$, $R_{21} (t, \xi)$, $P_{23} (t, \xi)$, $R_{32} (t, \xi)$ and $P_{31} (t, \xi)$. It is obvious from (41) that poles do not move

$$\zeta_{jk} (t) = \zeta_{jk} (0)$$

and we get the residue as

$$\gamma_K (t) = \gamma_K^{(o)} \exp (\sqrt{3} \zeta^{-1} \omega t) \tag{43}$$

Similarly we can calculate $\beta_K (t)$ and $\alpha_K (t)$.

The solution of (32) consists of three steps.

(*i*) Solve the spectral problem (33) at time $t = 0$ with **Q** $(x, 0)$ given by (35) and find the spectral data S.

(*ii*) Allow the spectral data to evolve to time t according to

$R_{13}(\xi, t)$, $P_{12}(t, \xi)$, $R_{21}(t, \xi)$, $P_{23}(t, \xi)$, $R_{32}(t, \xi)$

$P_{31}(t, \xi)$, $\zeta_{jk}(t)$, $\alpha_K(t)$, $\beta_K(t)$, $\gamma_K(t)$.

(*iii*) Invert the speetral transform to find Q (x, t), and hence $\underline{\sigma}(x, t)$ from the spectral data.

REFERENCES

1. Kblowitz M.J., Kaup D.J., Newell A.C., and Segur H., Studies in Appl. Maths. **53**, No. 4, 249 (1974).

2. Wadati, M., "Solitons", Topics in Current Physics 17, Bullough R.K., and Caudrey P.J., eds. (Springer-Verlag, Berlin, Heidelberg, New York 1980) PP. 287-299.

3. Fordy A., and Gibbons J., To be published (1982).

The Punjab University
Journal of Mathematics
Vol. XIV—XV (1981-82), pp. 155-164

COMMON FIXED POINTS OF SET-VALUED MAPPINGS

By
BRIAN FISHER

Department of Mathematics,
The University, Leicester,
LE1 7RH, U.K.

In the following, as in [1], we let (X, d) be a complete metric space and let $B(X)$ be the set of all nonempty, bounded subsets of X. The function $\delta(A, B)$ with A and B in $B(X)$ is defined by

$$\delta(A, B) = \sup \{d(a, b) : a \in A, b \in B\}.$$

If A consists of a single point a we write

$$\delta(A, B) = \delta(a, B).$$

If B also consists of a single point b we write

$$\delta(A, B) = \delta(a, b) = d(a, b).$$

It follows immediately that

$$\delta(A, B) = \delta(B, A) \geq 0,$$
$$\delta(A, B) \leq \delta(A, C) + \delta(C, B)$$

for all A, B and C in $B(X)$.

If now $\{A_n : n = 1, 2, \dots\}$ is a sequence of sets in $B(X)$, we say that it converges to the subset A of X if

(*i*) each point a in A is the limit of some convergent sequence $\{a_n\}$ with a_n in A_n for $n = 1, 2, \dots$,

(*ii*) for arbitrary $\varepsilon > 0$, there exists an integer N such that $A_n \subset A_\varepsilon$ for $n > N$, where A_ε is the union of all open spheres with centres in A and radius ε.

The set A is then said to be the limit of the sequence $\{A_n\}$.

The following lemma was proved in [1].

Lemma. If $\{A_n\}$ and $\{B_n\}$ are sequences of bounded subsets of a complete metric space (X, d) which converge to the bounded subsets A and B respectively, then the sequence $\{\delta(A_n, B_n)\}$ converges to $\delta(A, B)$.

Now let F be a mapping of a complete metric space (X, d) into $B(X)$. We say that the mapping F is continuous at a point x in X if whenever $\{x_n\}$ is a sequence of points in X converging to x, the sequence $\{Fx_n\}$ in $B(X)$ converges to Fx in $B(X)$. We say that F is a continuous mapping of X into $B(X)$ if F is continuous at each point x in X. We say that a point z in X is a fixed point of F if z is in Fz. If A is any nonempty subset of X we define the set FA by

$$FA = \bigcup_{a \in A} Fa.$$

We now prove the following theorem.

Theorem 1. Let F and G be continuous mappings of a complete metric space (X, d) into $B(X)$ satisfying the inequality

$$\delta(F^p x, G^p y) \leq c. \max\{\delta(F^r x, G^s y) : 0 \leq r, s \leq p\} \tag{1}$$

for all x, y in X, where $0 \leq c < 1$ and p is a fixed positive integer. If F and G also map $B(X)$ into itself, then F and G have a unique common fixed point z. Further z is the unique fixed point of F and G and $Fz = Gz = \{z\}$.

Proof. Since we are supposing that F and G map $B(X)$ into itself we note that both sides of inequallty (1) are finite. Further, if A and B are any sets in $B(X)$ then it follows easily that

$$\delta(F^p A, G^p B) \leq c. \max\{\delta(F^r A, G^s B) : 0 \leq r, s \leq p\} \tag{2}$$

both sides of the inequality again being finite.

Now let x be an arbitrary point in X and put $X_n = F^n x$ and $Y_n = G^n x$ for $n = 0, 1, 2, \ldots$, where $F^0 x = G^0 x = x$. Let us suppose that the set of real numbers

$$\{\delta(X_n, Y_p) : n = 1, 2, \ldots\}$$

is unbounded. Then there exists an integer $n > p$ such that

$$(1 - c)\,\delta(X_n, Y_p) > c.\max\{\delta(Y_s, Y_p) : 0 \leq s \leq p\}$$

and

$$\delta(X_n, Y_p) > \max\{\delta(X_r, Y_p) : 0 \leq r < n\}.$$

These inequalities imply that

$$\begin{aligned} c\,\delta(X_r, Y_s) &\leq c\,\delta(X_r, Y_p) + c\,\delta(Y_p, Y_s) \\ &< c\,\delta(X_n, Y_p) + (1 - c)\,\delta(X_n, Y_p) \\ &= \delta(X_n, Y_p) \end{aligned}$$

for $0 \leq r \leq n$ and $0 \leq s \leq p$. Thus

$$\delta(X_n, Y_p) > c.\max\{\delta(X_r, Y_s) : 0 \leq r \leq n;\ 0 \leq s \leq p\}.$$

However on using inequality (2) it follows that

$$\begin{aligned} \delta(X_n, Y_p) &\leq c.\max\{\delta(X_r, Y_s) : n - p \leq r \leq n;\ 0 \leq s \leq p\} \\ &< \delta(X_n, Y_p) \end{aligned}$$

from what we have just proved, giving a contradiction. This contradiction implies that

$$\sup\{\delta(X_n, Y_p) : n = 0, 1, 2, \ldots\} = M_1 < \infty.$$

Similarly we can prove that

$$\sup\{\delta(X_p, Y_n) : n = 0, 1, 2, \ldots\} = M_2 < \infty$$

and it follows that

$$\begin{aligned} &\sup\{\delta(X_r, Y_s) : r, s = 0, 1\ 2, \ldots\} \\ &\leq \sup\{\delta(X_r, Y_p) : r = 0, 1, 2, \ldots\} + \delta(Y_p, X_p) + \\ &\qquad + \sup\{\delta(X_p, Y_s) : s = 0, 1, 2, \ldots\} \\ &= M_1 + \delta(Y_p, X_p) + M_2 \\ &= M < \infty. \end{aligned}$$

Now for arbitrary $\varepsilon > 0$, choose a positive integer N such that $c^N M < \varepsilon$. Then if $m, n \geq N_p$ we have with repeated use of inequality (2)

$$\begin{aligned} \delta(X_m, Y_n) &\leq c.\max\{\delta(X_r, Y_s) : m - p \leq r \leq m;\ n - p \leq s \leq n\} \\ &\leq c^2.\max\{\delta(X_r, Y_s) : m - 2p \leq r \leq m;\ n - 2p \leq s \leq n\} \end{aligned}$$

$$\leq c^N . \max\{\delta(X_r, X_s) : m - N_p \leq r \leq m ; n - N_p \leq s \leq n\}$$
$$\leq c^N M < \varepsilon$$

and so

$$\delta(X_m, X_n) \leq \delta(X_m, Y_r) + \delta(Y_r, X_n) < 2\varepsilon$$

for $m, n, r \geq N_p$. Choosing a point x_n in X_n for $n = 1, 2, \ldots$ we have

$$d(x_m, x_n) \leq \delta(X_m, X_n) \leq 2\varepsilon$$

for $m, n > N_p$. The sequence $\{x_n\}$ is therefore a Cauchy sequence in the complete metric space X and so has a limit z in X. Further

$$\delta(z, Fx_n) \leq d(z, x_m) + \delta(x_m, Fx_n)$$
$$\leq d(z, x_m) + \delta(X_m, X_{n+1})$$

since x_m is in X_m and Fx_n is contained in X_{n+1} . Thus

$$\delta(z, Fx_n) < d(z, x_m) + \varepsilon$$

for $m, n+1 > N_p$. Letting m tend to infinity it follows that

$$\delta(z, Fx_n) \leq \varepsilon$$

for $n+1 > N_p$. Using the continuity of F and the lemma, it follows on letting n tend to infinity that

$$\delta(z, Fz) \leq \varepsilon .$$

Since ε is arbitrary $\delta(z, Fz) = 0$ and so we must have $Fz = \{z\}$.

We can prove similarly that there exists a point z' in X such that $Gz' = \{z'\}$. Then

$$d(z, z') = \delta(F^p z, G^p z')$$
$$\leq c. \max\{\delta(F^r z, G^s z') : 0 \leq r, s \leq p\}$$
$$= cd(z, z')$$

and so $z = z'$. The point z is therefore a common fixed point of F and G.

Now suppose that F has a second fixed point w so that w is contained in Fw and $F^n w$ is contained in $F^{n+1} w$ for $n = 1, 2, \ldots$, Then

$$\begin{aligned} d(w, z) &\leq \delta(F^p w, G^p z) \\ &\leq c.\ \max\{\delta(F^r w, G^s z) : 0 \leq r, s \leq p\} \\ &= c\,\delta(F^p w, G^p z) \end{aligned}$$

and it follows that

$$F^p w = \{z\} = \{w\}.$$

The point z is therefore the uniquc fixed point of F.

Similarly z is the unique fixed point of G. This completes the proof of the theorem.

Corollary. Let S and T be continuous mappings of a complete metric space (X, d) into itself satisfying the inequality

$$d(S^p x, T^p y) \leq c.\ \max\{d(S^r x, T^s y) : 0 \leq r, s \leq p\}$$

for all x, y in X, where $0 \leq c < 1$ and p is a fixed positive integer. Then S and T have a unique common fixed point z. Further z is the unique fixed point of S and T.

Proof. Define mappings F and G of X into B (X) by

$$Fx = \{Sx\},\ Gx = \{Tx\}$$

for all x in X. The conditions of the theorem are satisfied for F and G and so they have a unique common fixed point z. The point z is then the unique fixed point of S and T.

The result of this corollary was given in [2].

Theorem 2. Let F be a continuous mapping and G be a mapping of a complete metric space (X, d) into B (X) satisfying the inequality

$$\delta(F^p x, Gy) \leq c.\ \max\{F^r x, Gy), \delta(F^r x, y) : 0 \leq r \leq p\} \quad (3)$$

for all x, y in X, where $0 \leq c < 1$ and p is a fixed positive integer. If F also maps B (X) into itself, then F and G have a unique common fixed point z. Further z is the unique fixed point of F and G and $Fz = Gz = \{z\}$.

Proof. It follows from inequality (3) that

$$\delta(F^p A, Gy) \leq c.\ \max\{\delta(F^r A, Gy),\ \delta(F^r A, y) : 0 \leq r \leq p\}$$

for all A in B (X) and y in X.

Now let x be an arbitrary point in X and define X_n as in the proof of theorem 1. Letting y_0 be an arbitrary point in X we now define the sequence $\{y_n\}$ inductively. Having defined the point y_{n-1} we choose a point y_n in the set $Gy_{n-1} = Y_n$.

The assumption that the set of real numbers

$$\{ \delta (X_n , Y_1) : n = 1, 2, \ldots \}$$

is unbounded implies that there exists an integer $n > p$ such that

$$(1 - c) \ \delta (X_n , Y_1) > c. \max \{ \delta (y_0 , Y_1), \delta (Y_1 , Y_1) \}$$

and

$$\delta (X_n , Y_1) > \max \{\delta (X_r , Y_1) : 0 \leq r < n\}.$$

As in the proof of theorem 1, this leads to a contradiction so that

$$\sup \{ \delta (X_n , Y_1) : n = 0, 1, 2, \ldots \} = M_1 < \infty .$$

Similarly it follows that

$$\sup \{ \delta (X_p , Y_n) : n = 0, 1, 2, \ldots \} = M_2 < \infty$$

and so

$$\sup \{ \delta (X_r , Y_s) : r, s = 0, 1, 2, \ldots \} \leq M_1 + \delta (Y_1 , X_p) + M_2 = M < \infty .$$

Again for arbitrary $\varepsilon > 0$, choose a positive integer N such that $c^N M < \varepsilon$. Then if $m, n \geq N_p$ it follows that $\delta (X_m , Y_n) < \varepsilon$. As in the proof of theorem 1 it follows that F has a fixed point z and $Fz = \{z\}$. Also

$$d (y_m , y_n) \leq \delta (Y_m , Y_n)$$
$$\leq \delta (Y_m , X_r) + \delta (X_r , Y_n) < 2 \varepsilon$$

for $m, n, r \geq N_p$. The sequence $\{ y_n \}$ is therefore a Cauchy sequence in X and so has a limit z' in X. Further, if $m, n \geq N_p$

$$d (y_m , y_n) \leq \delta (y_m , Y_n) < 2 \varepsilon$$

and on letting m tend to infinity it follows that

$$\delta (z' , Y_n) \leq 2 \varepsilon$$

for $n > N_p$. Thus, on using inequality (3)

$$d(z, y_n) \leq \delta(F^p z, Gy_{n-1})$$
$$\leq c. \max \{ \delta(z, Gy_{n-1}), d(z, y_{n-1}) \}$$
$$\leq c. \max \{ \delta(z, Y_n), \delta(z, Y_{n-1}) \}$$
$$\leq c. \max \{ d(z, z') + \delta(z', Y_n), d(z, z') + \delta(z', Y_{n-1}) \}$$
$$\leq c. [d(z, z') + 2\varepsilon]$$

for $n + 1 > N_p$. On letting n tend to infinity it follows that

$$d(z, z') \leq c[d(z, z') + 2\varepsilon].$$

Since ε is arbitrary we must have $z = z'$.

We now have

$$\delta(z, Gz) = \delta(F^p z, Gz) \leq c. \delta(z, Gz)$$

and so $Gz = \{z\}$. The point z is therefore a common fixed point of F and G.

The uniqueness of z follows as in the proof of theorem 1. This completes the proof of the theorem.

The following corollary follows easily.

Corollary. Let S be a continuous mapping and T be a mapping of a complete metric space (X, d) into itself satisfylng the inequality

$$d(S^p x, Ty) \leq c. \max \{d(S^r x, Ty), d(S^r x, y) : 0 \leq r \leq p\}$$

for all x, y in X, where $0 \leq c < 1$ and p is a fixed positive integer. Then S and T have a unique common fixed point z. Further z is the unique fixed point of S and T.

The next theorem also holds and was proved in [3].

Theorem 3. Let F and G be mappings of a complete metric space (X, d) into B (X) satisfying the inequality

$$\delta(Fx, Gy) \leq c. \max \{ \delta(Fx, y), \delta(x, Gy), d(x, y) \}$$

for all x, y in X, where $0 \leq c < 1$. Then F and G have a unique common fixed point z. Further z is the unique fixed point of F and G and $Fz = Gz = \{z\}$.

We finally prove a theorem for compact metric spaces.

Theorem 4. Let F and G be continuous mappings of a compact metric space (X, d) into B (X) satisfying the inequality

$$\delta(F^p x, G^p y) < \max\{\delta(F^r x, G^s y) : 0 \leq r, s \leq p\} \tag{4}$$

for all x, y in X for which the right-hand side of the inequality is positive, where p is a fixed positive integer. Then F and G have a unique common fixed point z. Further z is the unique fixed point of F and G and $Fz = Gz = \{z\}$.

Proof. We note first of all that since X is compact every subset of X is bounded and so both sides of inequality (4) are finite.

Let us suppose first of all that the right-hand side of inequality (4) is positive for all x, y in X. Define the real-valued function $f(x, y)$ on X^2 by

$$f(x, y) = \frac{\delta(F^p x, G^p y)}{\max\{\delta(F^r x, G^s y) : 0 \leq r, s \leq p\}}.$$

Then if $\{(x_n, y_n)\}$ is an arbitrary sequence in X^2 converging to (x, y), it follows easily from the lemma and the continuity of F and G that the sequence $\{f(x_n, y_n)\}$ converges to $f(x, y)$. The function f is therefore a continuous function defined on the compact metric space X^2 and so achieves its maximum value c. Inequality (4) implies that $c < 1$ and it follows that the conditions of theorem 1 are satisfied. Hence F and G have a unique common fixed point z and $Fz = Gz = \{z\}$.

Now let us suppose that the right-hand side of inequality (4) is zero for some x, y in X. Then

$$Fx = Gy = \{x\} = \{y\}$$

is a singleton z and it follows that z is a common fixed point of F and G and $Fz = Gz = \{z\}$.

Finally let us suppose that F has a second fixed point w. If $Fw \neq \{w\}$ then inequality (4) holds and we have

$$\begin{aligned}\delta(F^p w, G^p z) &< \max\{\delta(F^r w, G^s z) : 0 \leq r, s \leq p\}\\ &= \delta(F^p w, G^p z)\end{aligned}$$

since $F^{n-1} w$ is contained in $F^n w$ for $n = 1, 2, \ldots$. This gives a contradiction and so $Fw = \{w\}$.

If $z \neq w$ then inequality (4) holds and so

$$d(w, z) = \delta(F^p w, G^p z) < d(w, z),$$

giving a contradiction. The fixed point z of F must therefore be unique.

Similarly z is the unique fixed point of G. This completes the proof of the theorem.

The corollary follows easily.

Corollary. Let S and T be continuous mappings of a compact metric space (X, d) into itself satisfying the inequality

$$d(S^p x, T^p y) < \max\{d(S^r x, T^s y) : 0 \leq r, s \leq p\}$$

for all x, y in X for which the right-hand side of the inequality is positive, where p is a fixed positive integer. Then S and T have a unique common fixed point z. Further z is the unique fixed point of S and T.

REFERENCES

1. B. Fisher, "Common fixed points of mappings and set-valued mappings", Rostock. Math. Kolloq., 69-77, **18** (1981).
2. B. Fisher, "Results on common fixed points on complete metric spaces", Glasgow Math. J., 165-7, **21** (1980).
3. B. Fisher, "Common fixed points for set-valued mappings", Indian J. Math., to appear.

The Punjab University.
Journal of Mathematics
Vol, XIV—XV (1981-82), pp. 165-174

ON THE WARING FORMULA FOR THE POWER SUMS

By

P.G. TODOROV

Department of Mathematics,
Paissii Hil ndarski University,
4000 Plovdiv, Bulgaria.

Abstract :

In this papar is given a new form of the Waring formula which is convenient for applications. Thus, in particular, for the first time the power sums of the cubic equation are found in elosed form.

Let $n \geq 2$ be an integer and let

$$\sigma_k = C_n^k \left[x_1 , x_2 , \ldots , x_n \right], \qquad (1 \leqq k \leqq n) \quad (1)$$

be the elementary symmetric polynomials where the symbol on the right-hand side denotes the sum of products of the variables $x_1 , \ldots , x_n$ taken as the k-th combinations. If $x_1 , \ldots , x_n$ are roots of the algebraic equation

$$x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n = 0, \qquad (2)$$

then the values of the symmetric polynomials (1) are equal to

$$\sigma_k = (-l)^k a_k , \qquad (1 \leqq k \leqq n). \quad (3)$$

Let

$$S_m (n) = x_1{}^m + x_2{}^m + \ldots + x_n{}^m \qquad (m = 1, 2, \ldots) \quad (4)$$

denote the sums of the m-th powers (abbreviated the "power sums") of the roots of equation (2), An expression of $S_m(n)$ in terms of a_k (or σ_k) yields the Waring formula

$$S_m(n) = m \sum (-1)^{\nu_1 + m + \nu_n} \frac{(\nu_1 + m + n -)!}{\nu_1! \dots \nu_n!} a_1^{\nu_1} \dots a_n^{\nu_n} \quad (5)$$

where the sum is taken over all non-negative integers $\nu_1, \dots, \nu_n$ for which

$$\nu_1 + 2\nu_2 + \dots + n\nu_n = m, \quad (6)$$

(See Ch. Jordan [1], *p.* 595, Comtet [2], pp, 158-159, point 9, example (3) and p. 140 and our paper [3], pp. 82-85, section 2).

The Waring formula is cumbersome and has no applications. Therefore, it has not been observed so far that even for the simplest case $n = 2$ the formula (5) can solve the problem altogether. In fact from (6) we obtain $\nu_1 = m - 2k$, $k ! = \nu_2$. Since $\nu_1 \geq 0$, it follows that $k \leqq m/2$, i.e. $0 \leqq k \leqq [m/2]$, where, here and elsewhere in this paper, [X] for an arbitrary X denotes the greatest integer in X. Thus from (5) we obtain immediately the Newton classic formula

$$S_m(2) = x_1^m + x_2^m = m \sum_{k=0}^{[m/2]} \frac{(-1)^{m-k}}{m-k} \binom{m-k}{k} a_1^{m-2k} a_2^k \quad (7)$$

$$(m \geq 1)$$

for the power sums of the roots x_1, x_2 of the quadratic equation

$$x^2 + a_1 x + a_2 = 0, \quad (8)$$

The above proof of Formula (7) is simpler than the well-known proofs of the same formula (compare, for example, Comtet [2], pp. 155-156, point 1).

For $n = 3, 4, \dots$, and an arbitrary positive integer m, the solution of the equation (6) depends on 2, 3 etc. parameters and, therefore, Formula (5) has no simple form. Thus, the problem arises whether it is possible to reduce the number of the parameters so that we can obtain simple formulas for $S_m(n)$ at least for the initial values $n = 3, 4, \ldots$, and an arbitrary positive integer m. It is shown that this is possible if, instead of the Waring formula (5–6), we offer the following formula :

Theorem 1. For arbitrary positive integers $m \geq 1$ and $n \geq 2$ we have

$$S_m(n) = m \sum_{k=0}^{m-1} (-1)^{m-k} (m-k-1)!\, B_{m,\, m-k}(a_1, \ldots, a_n) \qquad (9)$$

where $B_{m,\, m-k}(a_n, \ldots, a_n)$ are the homogeneous isobaric polynomials

$$B_{m,\, m-k}(a_1, \ldots, a_n) = \sum \frac{a_1^{\nu_1} \ldots, a_n^{\nu_n}}{\nu_1! \ldots, \nu_n!} \qquad (10)$$

of degree $m-k$ and of weight m in the variables $a_1, \ldots, a_n$, i.e. the sum in (10) is taken over all non-negative integers $\nu_1, \ldots, \nu_n$ satisfying

$$\nu_1 + \nu_2 + \ldots + \nu_n = m - k, \quad \nu_1 + 2\nu_2 + \ldots + n\,\nu_n = m. \qquad (11)$$

Proof. From (6) we obtain

$$1 \leqq \nu_1 + \nu_2 + \ldots +, \nu_n \leqq m, \qquad (12)$$

It follows from (12) that we can also join to equation (6) the first equation of (11) for $k=0, 1, \ldots, m-1$. Hence the Waring formula (5) can be written in the form (9). This completes the proof of Theorem 1.

APPLICATIONS OF THEOREM 1

For $n = 2$ we immediately obtain the Newton classic formula (7) again. For $n = 3$ we obtain the following new result :

Theorem 2. Let

$$-a_1 = x_1 + x_2 + x_3, \quad a_2 = x_1 x_2 + x_1 x_3 + x_2 x_3, \quad -a_3 = x_1 x_2 x_3 \qquad (13)$$

be the elementary symmetric polynomials of the roots x_1, x_2, x_3 of the cubic equation

$$x^3 + a_1 x^2 + a_2 x + a_3 = 0. \tag{14}$$

Then for $m = 1, 2, 3, \dots$, the power sums

$$S_m(3) = x_1^m + x_2^m + x_3^m = m \sum_{k=0}^{[2m/3]} (-1)^{m-k} (m-k-1)!$$

$$\cdot \sum_{\nu=s}^{[k/2]} \frac{a_1^{m-2k+\nu} \, a_2^{k-2\nu} \, a_3^{\nu}}{(m-2k+\nu)! \, (k-2\nu)! \, \nu!} \tag{15}$$

where $s = 0$ for $0 \leqq k \leqq m/2$ and $s = 2k-m$ for $m/2 \leqq k \leqq 2m/3$.

Proof. For $n = 3$ from (11) we obtain

$$\gamma_1 = m - 2k + \gamma \geqq 0, \ \nu_2 = k - 2\nu \geqq 0 \qquad (\nu\,! = \nu_3 \geqq 0), \tag{16}$$

It follows from (16) that $\nu \geqq 2k - m$ and $\nu \leqq k/2$. This is possible if $k \leqq 2m/3$. Hence

$$\begin{cases} 0 \leqq \nu \leqq \dfrac{k}{2} \text{ for } 0 \leqq k \leqq \dfrac{m}{2}, \\ 2k - m \leqq \nu \leqq \dfrac{k}{2} \text{ for } \dfrac{m}{2} \leqq k \leqq \dfrac{2m}{3}. \end{cases} \tag{17}$$

Thus for $n = 3$ the formula (9-11) by means of (16-17) takes the form (15). This completes the proof of Theorem 2.

We note two important corollaries of Theorem 2.

If in (13-14) we set $a_1 = 0$, $a_2 = p$, $a_3 = q$, then in the inner sum of (15) only for $\nu = 2k - m$ the terms do not vanish. This is possible for $m/2 \leqq k \leqq 2m/3$. Thus we obtain the following new formula :

Corollary 1. x_1, x_2, x_3 are the roots of the cubic equation

$$x^3 + px + q = 0, \tag{18}$$

then for $m = 1, 2, 3, \dots$ the power sums

$$S_m(3) = x_1^m + x_2^m + x_3^m \tag{19}$$

$$= m. \sum_{m/2 \leqq k \leqq 2m/3} \frac{(-1)^{m-k}}{m-k} \binom{m-k}{2k-m} p^{2m-3k} q^{2k-m}$$

Where the summation is taken over the integers k.

If in (13-14) we get $a_1 = p$, $a_2 = 0$, $a_3 = q$, then in the inner sum of (15) only for $\nu = k/2$ the term does not vanish. Hence k should be an even number. If we substitute k for $2k$, $0 \leqq k \leqq m/3$ in (15), then we shall obtain another new formula :

Corollary 2. x_1, x_2, x_3 are the roots of the cubic equation

$$x^3 + px^2 + q = 0, \tag{20}$$

then for $m = 1, 2, 3, \ldots$ the power sums

$$S_m(3) = x_1^m + x_2^m + x_3^m = (-1)^m\, m \sum_{k=0}^{[m/3]} \binom{m-2k}{k} \frac{p^{m-3k}\, q^k}{m-2k}. \tag{21}$$

By substituting $x = 1/y$ it is clear that the formulas (19) and (21) express the power sums of the equations (20) and (18), respectively, when m is a negative integer.

Further, for $n = 4$, the solutions of the system (11) depend on two parameters ν_3 and ν_4 and according to our method the polynomial (10) will be expressed by a double sum. Thus, we can obtain the corresponding formula for $S_m(4)$ of the general equation of the fourth degree $\left[\,2\,\right]$ $(n = 4)$. Here the problem is reduced if some coefficient vanishes. For example, if $a_1 = 0$, this yields $\nu_1 = 0$ in (11). Hence the solutions of the system (11) will depend only on the parameter ν_4 and the polynomial (10) will be expressed by a simple sum similar to Theorem 2. It is evident that our method can be applied also for $n = 5, 6, \ldots$ and for an arbitrary positive integer m. It is clear that the corresponding formulas must be constructed only if this is necessary.

So far we have considered the applications of Theorem 1 for given $n = 2, 3, 4, \ldots$ and an arbitrary positive integer m. Conversely, for given $m = 1, 2, 3, \ldots$ and a fixed positive integer $n \geqq 2$, for abbreviation of the calculation, it is better to use the following still more precise formula :

Theorem 3. I. If $1 \leqq m \leqq n$, $n \geqq 2$, then the power sum

$$S_m(n) = m \sum_{k=0}^{m-1} (-1)^{m-k} (m-k-1)!\, B_{m,\, m-k} \tag{22}$$

where the polynomial

$$B_{m,\, m-k} = \sum \frac{a_1^{\nu_1} \dots a_{k+1}^{\nu_{k+1}}}{\nu_1! \dots \nu_{k+1}!}, \tag{23}$$

i.e. the sum is taken over all non-negative integers $\nu_1, \dots, \nu_{k+1}$ satisfying

$$\nu_1 + \nu_2 + \dots + \nu_{k+1} = m-k,\ \nu_1 + 2\nu_2 + \dots + (k-1)\ \nu_{+1} = m, \tag{24}$$

II. If $m \geq n$, $n \geqq 2$, then the power sum

$$S_m(n) = m \left(\sum_{k=0}^{n-1} + \sum_{k=n}^{m-1} \right) (-1)^{m-k} (m-k-1)!\, B_{m,\, m-k} \tag{25}$$

where for $0 \leqq k \leqq n-1$ the polynomial

$$B_{m,\, m-k} = \sum \frac{a_1^{\nu_1} \dots a_{k+1}^{\nu_{k+1}}}{\nu_1! \dots \nu_{k+1}!} \tag{26}$$

and the sum is taken over all non-negative integers $\nu_1, \dots, \nu_{k+1}$ satisfying

$$\nu_1 + \nu_2 + \dots + \nu_{k+l} = m-k,\ \nu_1 + 2\nu_2 + \dots + (k+1)\ \nu_{k+1} = m, \tag{27}$$

and for $n \leqq k \leqq m-1$ the polynomial

$$B_{m,\, m-k} = \frac{a_1^{\nu_1} \dots a_n^{\nu_n}}{\nu_1! \dots \nu_n!} \tag{28}$$

where the sum is taken over all non-negative integers $\nu_1, \dots, \nu_n$ satisfying

$$\nu_1 + \nu_2 + \dots + \nu_n = m=k,\ \nu_1 + 2\nu_2 + \dots + n\,\nu_n = m. \tag{29}$$

Proof. If we substract the first equation of (11) from the second one, then we shall obtain the subsidiary equation

$$\sum_{s=2}^{n} (s-1)\, \nu_s = k \qquad (0 \leqq k \leqq m-1). \tag{30}$$

Hence if $\nu_s \geq 1$ for $s > k + 1$, then $(s-1)\ \nu_s \geq s - 1 > k$ and the equation (30) has no meaning. Therefore $\nu_s = 0$ for $s > k + 1$. Now, if we consider separately the two cases $1 \leqq m \leqq n$ and $m \geq n$, then we conclude that Theorem 1 can be modified into Theorem 3. This completes our proof.

Now we shall yield an effective method for solving the system (24). For $k = 0$ we obtain $\nu_1 = m$ and the polynomial (23) is equal to

$$B_{m,\, m} = \frac{a_1{}^m}{m\,!}\ , \qquad (m \geq 1). \qquad (31)$$

For $1 \leqq k \leqq m - 1$, $m \geq 2$ the number of the unknowns in the system (24) can be reduced by one. In fact, here the subsidiary equation (30) is

$$\nu_2 + 2\nu_3 + \dots + (k-1)\ \nu_k + k\nu_{k+1} = k, \qquad (32)$$

from which we obtain $k\nu_{k+1} \leqq k$, i.e. for ν_{k+1} the values 1 and 0 are possible. For $\nu_{k+1} = 1$ from (32) it follows that $\nu_k = \dots = \nu_2 = 0$, $k \geq 2$, whence and from (24) we find $\nu_1 = m - k - 1$. Hence the system (24) always has the solution

$$\nu_1 = m-k-1,\ \nu_2 = \dots = \nu_k = 0,\ \nu_{k+1} = 1 \quad (2 \leqq k \leqq m-1,\ m \geq 3), \ (33)$$
$$\nu_1 = m - 2,\ \nu_2 = 1 \qquad (k = 1,\ m \geqq 2).$$

For $\nu_{k+1} = 0$ from (24) and (32) we obtain the system

$$\nu_1 + \nu_2 + \dots + \nu_k = m - k,\ \nu_1 + 2\nu_2 + \dots + k\nu_k = m \qquad (34)$$
$$(2 \leqq k \leqq m - 1,\ m \geqq 3)$$

and the subsidiary equation

$$\nu_2 + 2\nu_3 + \dots + (k-1)\ \nu_k = k \quad (2 \leqq k \leqq m-1,\ m \geqq 3) \qquad (35)$$

which determine the solutions of the system (24) having the form $(\nu_1, \dots, \nu_k, 0)$. Further, from (35) we obtain $(k-1)\ \nu_k \leqq k$. By this inequality it follows that the possible values of ν_k are the non-negative integers from the closed interval $\left[\ 0,\ k\ /\ (k-1)\ \right]$, i.e. the number of the unknowns in (34) and (35) again is reduced by one. Thus, by successive elimination of the unknowns we obtain that $\nu_{k-1}\ \varepsilon$

$\left[0 , k / (k-2) \right]$, ν_{k-2} ε $\left[0 , k / (k-3) \right]$, ..., ν_3 ε $\left[0, k/2 \right]$,

ν_1 ε $\left[0, m-k \right]$.

Hence by elimination and verification it is easy to find all solutions of the system (34) successive for $k = 2, 3, \dots, m-1$, $m \geqq 3$.

By this method we can calculate the polynomial (23) in the general case. Thus for $k = 1$ from (33) and (23) we obtain

$$B_{m,\, m-1} = \frac{a_1{}^{m-2}\, a_2}{(m-2)\,!}, \qquad (m \geqq 2). \quad (36)$$

For $k = 2$ from (33) we obtain $\nu_1 = m-3$, $\nu_2 = 0$, $\nu_3 = 1$ and from (34) we find $\nu_1 = m-4$, $\nu_2 = 2$ ($\nu_3 = 0$). Hence the polynomial (23) has the form

$$B_{m,\, m-2} = \frac{a_1{}^{m-3}\, a_3}{(m-3)\,!} + \frac{a_1{}^{m-4}\, a_2{}^2}{(m-4)\,!\, 2\,!} \qquad (m \geqq 4). \quad (37)$$

For $k = 3$ from (33) we have the solution $\nu_1 = m-4$, $\nu_2 = \nu_3 = 0$, $\nu_4 = 1$ and from (34) and (35) by our algorithm we easily obtain two other solutions $\nu_1 = m-5$, $\nu_2 = 1$, ($\nu_4 = 0$) and $\nu_1 = m-6$, $\nu_2 = 3$, $\nu_3 = 0$, ($\nu_4 = 0$). Hence the polynomial (23) has the form

$$B_{m,\, m-3} = \frac{a_1{}^{m-4}\, a_4}{(m-4)\,!} + \frac{a_1{}^{m-5}\, a_2\, a_3}{(m-5)\,!} + \frac{a_1{}^{m-6}\, a_2{}^3}{(m-6)\,!\, 3!} \quad (m \geqq 6). \quad (38)$$

It is clear that we can continue in this way also for $k=4, 5, \dots, m-1$ and if m is a given positive integer, we calculate all polynomials (23). The last polynomial (23) ($k = m-1$) is

$$B_{m\, ,\, 1} = a_m\, , \qquad (m \geqq 1). \quad (39)$$

For a given m the system (27) and (29) can be solved by the same method setting successively $n = 2, 3, \dots, m-1$.

In conclusion we note that in the application of Theorem 1 or Theorem 3 to a concrete equation (2), the calculations are considerably reduced if one or several coefficients a_j vanish. Then we get $\nu_j = 0$ in (11) or in (24) ((34) and (35), respectively), (27) and (29).

REFERENCES

1. Charles Jordan, Calculus of Finite Differences, Chelsea, New York, 1965 (reprint 1979).
2. L. Comtet, Advanced Combinatorics. The Art of Finite and Infinite Expansion, Reidel, Dordrecht and Boston, Mass., 1974.
3. P.G. Todorov, New Explicit Formulas for the Coefficients of *p*-Symmetric Functions, Proc. Amer. Math. Soc., **77** (1979), pp. 81-86.

The Punjab University
Journal of Mathematics
Vol. XIV—XV (1981-82), pp. 175-182

A NOTE ON CERTAIN CLASSES OF ANALYTIC FUNCTIONS

By

SHIGEYOSHI OWA

Department of Mathematics
Kinki University, Osaka, Japan

Abstract.

Recently, M.S. Ganesan studied the class $P_\alpha(A, B)$ of analytic functions in the unit disk. In this paper, we shall define the subclass $\tilde{P}_\alpha(A, B)$ of $P_\alpha(A, B)$ and a class $\tilde{R}_\alpha(A, B)$ of analytic functions in the unit disk. And the object of the present paper is to show the interesting coefficient estimates for the classes $\tilde{P}_\alpha(A, B)$, $\tilde{R}_\alpha(A, B)$, $\tilde{P}_\alpha(A, 0)$ and $\tilde{R}_\alpha(A, 0)$.

1. Introduction.

Let $P_\alpha(A, B)$ denote the class of functions

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n$$

which are analytic in the unit disk $U = \{z : |z| < 1\}$ and satisfying the condition

$$\frac{f(z)}{g(z)} = \frac{1 + Aw(z)}{1 + Bw(z)} \qquad (z \in U),$$

AMS (MOS) subject classifications (1980). 30 C 45.

where $-1 \leqq B < A \leqq 1$ and $w(z)$ is analytic in the unit disk U with $w(0) = 0$ and $|w(z)| < 1$ for $z \in U$. Here $g(z)$ is a starlike function of order α $(0 \leq \alpha < 1)$ with respect to the origin in the unit disk U, that is,

$$\operatorname{Re}\left(\frac{zg'(z)}{g(z)}\right) > \alpha \qquad (z \in U).$$

M.S. Ganesan [1] has obtained a distortion theorem, the coefficient estimates and a radius of starlikeness for this class. On taking $A = \beta$ and $B = -\lambda\beta$ with $w(z)$ replaced by $-w(z)$, this class becomes the class $S_\lambda(\alpha, \beta)$ studied by R.M. Goel and N.S. Sohi [2].

Let $T^*(\alpha)$ and $C(\alpha)$ denote the classes of functions

$$f(z) = z - \sum_{n=2}^{\infty} a_n z^n \qquad (a_n \geqq 0)$$

which are analytic and starlike of order α $(0 \leqq \alpha < 1)$ with respect to the origin in the unit disk U and which are analytic and convex of order α $(0 \leqq \alpha < 1)$ in the unit disk U, respectively.

For these classes, H. Silverman [3] showed the following lemmas.

Lemma 1. A function

$$f(z) = z - \sum_{n=2}^{\infty} a_n z^n \qquad (a_n \geqq 0)$$

is in the class $T^*(\alpha)$ if and only if

$$\sum_{u=0}^{\infty} (n - \alpha) a_n \leqq 1 - \alpha .$$

Lemma 2. A function

$$f(z) = z - \sum_{n=2}^{\infty} a_n z^n \qquad (a_n \geqq 0)$$

is in the class C (α) if and only if

$$\sum_{n=2}^{\infty} n(n-\alpha)\, a_n \leqq 1-\alpha.$$

Let $\tilde{P}_{\alpha}$ (A, B) denote the class of functions

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n$$

which are analytic in the unit disk U and satisfying the condition

$$\frac{f(z)}{g(z)} = \frac{1 + Aw(z)}{1 + Bw(z)} \qquad (z \in U). \qquad (1)$$

where $-1 \leqq B < A \leqq 1$, $w(z)$ is analytic in the unit disk U with $w(0) = 0$ and $|w(z)| < 1$ for $z \in U$ and

$$g(z) = z - \sum_{n=2}^{\infty} b_n z^n \qquad (b_n \geqq 0)$$

is in the class T* (α) ($0 \leqq \alpha < 1$). Further let $\tilde{R}_{\alpha}$ (A, B) denote the class of functions

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n$$

which are analytic in the unit disk U and satisfying the condition (1) for $g(z) \in C(\alpha)$.

2. Coefficient Estimates for $\tilde{P}_{\alpha}$ (A, B) and $\tilde{R}_{\alpha}$ (A, B).

Theorem 1. Let a function

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n$$

be in the class $\tilde{P}_{\alpha}$ (A, B). Then we have

$$| a_2 | \leqq (A - B) + \frac{1 - \alpha}{2 - \alpha} .$$

The estimate is sharp. Further, for $B \geqq 0$, we have

$$| a_3 | \leqq (1 + B)(A - B) + \frac{1 - \alpha}{2 - \alpha} (A + B) + \frac{1 - \alpha}{3 - \alpha} .$$

Proof. We employ the same technique as used by M.S. Ganesan [1].

Let

$$\frac{f(z)}{g(z)} = \frac{1 + Aw(z)}{1 + Bw(z)}$$

and

$$g(z) = z - \sum_{n=2}^{\infty} b_n z^n \qquad (b_n \geqq 0).$$

Then we have

$$w(z) = \frac{f(z) - g(z)}{Ag(z) - Bf(z)} \tag{2}$$

Putting

$$w(z) = \sum_{n=1}^{\infty} c_n z^n$$

and substituting the power series for $f(z)$, $g(z)$ and $w(z)$ in (2), we obtain

$$\left\{ A\left(z - \sum_{n=2}^{\infty} b_n z^n \right) - B\left(z + \sum_{n=2}^{\infty} a_n z^n \right) \right\} \left(\sum_{n=1}^{\infty} c_n z^n \right)$$

$$= \sum_{n=2}^{\infty} (a_n + b_n) z^n . \tag{3}$$

Equating the coefficients of z^2 and z^3 on both sides of (3) we get

$$a_2 = (A - B)\, c_1 - b_2 \tag{4}$$

and

$$a_3 = (A - B)\, c_2 - (Ab_2 + Ba_2)\, c_1 - b_3 . \tag{5}$$

Since $|c_1| \leqq 1$ and $|c_2| \leqq 1 - |c_1|^2$ and Lemma 1 implies that $b_2 \leqq (1 - \alpha) / (2 - \alpha)$ and $b_3 \leqq (1 - \alpha) / (3 - \alpha)$, (4) and (5) give the required estimates. Further the estimate for $|a_2|$ is sharp for

$$\frac{f(z)}{g(z)} = \frac{1 - Az}{1 - Bz}$$

and

$$g(z) = z - \frac{1 - \alpha}{2 - \alpha} z^2 .$$

Theorem 2. Let a function

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n$$

be in the class $\tilde{R}_\alpha$ (A, B). Then we have

$$|a_2| \leqq (A - B) + \frac{1 - \alpha}{2(2 - \alpha)} ,$$

The estimate is sharp for

$$\frac{f(z)}{g(z)} = \frac{1 - Az}{1 - Bz}$$

and

$$g(z) = z - \frac{1 - \alpha}{2(2 - \alpha)} z^2 .$$

Further for $B \geqq 0$, we have

$$|a_3| \leqq (1 + B)(A - B) + \frac{1 - \alpha}{2(2 - \alpha)} (A + B) + \frac{1 - \alpha}{3(3 - \delta)} .$$

The proof of Theorem 2 is obtained by using the same technique sa in the proof of Theorem 1 with the aid of Lemma 2.

Remark 1. We have not been able to obtain sharp estimates for $|a_n|$ $(n \geqq 3)$ for the classes $\tilde{P}_\alpha(A, B)$ and $\tilde{R}_\alpha(A, B)$.

3. Coefficient estimates for $\tilde{P}_\alpha(A, B)$ and $\tilde{R}_\alpha(A, B)$.

Theorem 3. Let a function

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n$$

be in the class $\tilde{P}_\alpha(A, 0)$. Then we have

$$|a_n| \leqq \frac{A(3-2\alpha)}{2-\alpha} + \frac{1-\alpha}{n-\alpha}$$

for $n \geqq 2$.

Proof, Since $f(z)$ belongs to the class $\tilde{P}_\alpha(A, 0)$, we have

$$\frac{f(z)}{g(z)} = 1 + Aw(z), \tag{6}$$

where $w(z)$ is an analytic function in the unit disk U satisfies the conditions $w(0) = 0$ and $|w(z)| < 1$ for $z \in U$. Let the function $w(z)$ have the expansion

$$w(z) = \sum_{n=1}^{\infty} c_n z^n \qquad (z \in U).$$

Then on substituting the power series for functions $f(z)$, $g(z)$ and $w(z)$ in (6), we get

$$\sum_{n=2}^{\infty} (a_n + b_n) z^n$$

$$= A\left(z - \sum_{n=2}^{\infty} b_n z^n\right)\left(\sum_{n=1}^{\infty} c_n z^n\right). \tag{7}$$

Equating the coefficients of z^n on both sides of (7),

$$a_n + b_n = A\left(c_{n-1} - \sum_{n-2}^{n-1} b_m c_{n-m}\right).$$

Hence we obtain

$$|a_n| \leqq b_n + A\left(|c_{n-1}| + \sum_{m=2}^{n-1} b_m |c_{n-m}|\right).$$

Since $|w(z)| < 1$ for $z \in U$, $|c_n| \leqq 1$ for $n = 1, 2, 3, \ldots$, and hence

$$|a_n| \leqq b_n + A\left(1 + \sum_{m=2}^{n-1} b_m\right).$$

Further Lemma 1 implies that

$$\sum_{m=2}^{n-1} b_m \leqq \frac{1-\alpha}{2-\alpha}$$

and $b_n \leqq (1-\alpha)/(n+\alpha)$. This gives the required estimate.

Theorem 4. Let a function

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n$$

be in the class $\tilde{R}_\alpha(A, 0)$. Then we have

$$|a_n| \leqq \frac{A(5-3\alpha)}{2(2-\alpha)} + \frac{1-\alpha}{n(n-\alpha)}$$

for $n \geqq 2$.

The proof of Theorem 4 is obtained by using the same technique as in the proof of Theorem 3 in conjunction with Lemma 2.

Remark 2. We have not been able to obtain sharp estimates for $|a_n|$ in Theorem 3 and Theorem 4.

ACKNOWLEDGEMENT

The author thankfully acknowledges the kind and helpful guidance of Dr. M.H. Kazi, in preparation of this paper.

REFERENCES

1. M.S. Ganesan : On certain classes of analytic functions, Indian J. Pure Appl. Math., **13** (1982), 47-57.

2. R.M. Goel and N.S. Sohi : On certain classes of analytic functions, Indian J. Pure Appl. Math., **11** (1980), 1308-1324.

3. H. Silverman : Univalent functions with negative coefficients, Proc. Amer. Math. Soc., **51** (1975), 109-116.

The Punjab University
Journal of Mathematics
Vol. XIV—XV (1981-82), pp. 183-192

A SUBCLASS OF CLOSE-TO-CONVEX FUNCTIONS

By

K. INAYAT NOOR AND NAEELA AL-DIHAN

Mathematics Department,
Science College of Education for Girls
Sitteen Road, Malaz, Riyadh, Saudi Arabia.

Abstract.

A new subclass of close-to-convex functions is defined, and its relationship with other subclasses, coefficient problem, distortion theorems, arclength problem and other properties are studied.

Let A denote the class of functions $f : f(z) = z + \sum_{n=2}^{\infty} a_n z^n$, which are analytic in the unit disc $E = [z : |z| < 1]$. By S, K, S* and C, denote the subclasses of A, which are respectively univalent, close-to-convex, starlike and convex in E. In [9] a subclass C* of univalent functions was introduced and studied. A function f belongs to C* if and only if there exists a convex function g such that, for $z \varepsilon E$

$$\text{Re}\ \frac{(z f'(z))'}{g'(z)} > 0$$

$f \varepsilon C^*$ is called quasi-convex. It is shown [9] that $f \varepsilon C^*$ if and only if and only if $z f' \varepsilon K$ and $C \subset C^* \subset K \subset S$.

We now have the following :

Definition 1. Let $f : f(z) = z + \sum_{n=2}^{\infty} a_n z^n$ be analytic in E.

Then $f \varepsilon K^*$ if and only if there exists a starlike function g such that, for $z \varepsilon E$,

$$\operatorname{Re} \frac{(z f'(z))'}{g'(z)} > 0 \tag{1}$$

Theorem 1.

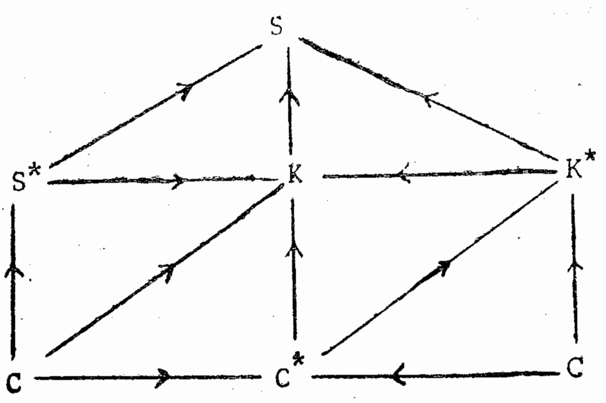
All the functions $f \varepsilon K^*$ are close-to-convex and hence univalent.

Proof. The proof follows immediately by using a well-known lemma due to Libera [2]. We have, for $g \varepsilon S^*$ and $z \varepsilon E$,

$$\operatorname{Re} \frac{(z f'(z))'}{g'(z)} > 0 \text{ implies } \operatorname{Re} \frac{z f'(z)}{g(z)} > 0.$$

Hence $f \varepsilon K^*$ is close-to-convex and since all close-to-convex functions are univalent [1], $f \varepsilon K^*$ is also univalent.

Remark 1. From definition 1, it is clear that $C \subset C^* \subset K^* \subset K \subset S$ and it is known [5] that $C \subset S^* \subset K \subset S$. We can thus write



where arrows indicate set inclusion.

Furthermore the function f_* given by $f_*(z) = \frac{1-i}{2} \frac{z}{1-z} - \frac{1+i}{2} \log(1-z)$ belongs to C^*, see [7], and since $C^* \subset K^*$ and so $f_* \varepsilon K^*$, but for ε sufficiently small, $\operatorname{Re} \frac{z f_*'(z)}{f_*(z)} < 0$, where $z = e^{i\theta}$, $-\varepsilon < \theta < 0$. This implies $f_* \notin S^*$.

We now proceed to give coefficient result and distortion theorems.

Theorem 2.

Let $f \in K^*$ and be given by $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$. Then

$$| a_n | \le \frac{2n^2+1}{3n}, \text{ for all } n.$$

This result is sharp.

Proof. Since $f \in K^*$, so by definition, there exists a $g \in S^*$ such that for $z \in E$,

$$(z f'(z))' = g'(z) h(z), \quad \text{Re } h(z) > 0.$$

Let $g(z) = z + \sum_{n=2}^{\infty} b_n z^n$, and $h(z) = 1 + \sum_{n=1}^{\infty} c_n z^n$. Then

$$n^2 a_n = [c_{n-1} + 2b_2 c_{n-2} + \ldots\ldots + (n-1) b_{n-1} c_1] + nb_n$$

Now it is well-known [5] that $| c_n | \le 2$ and $| b_n | \le n$, for all n. So we have

$$| a_n | \le \frac{2n^2+1}{3n}, \text{ for all } n.$$

The function f_0, is taken with respect to the Koebe function, belongs to K^*, and defined by

$$f_0(z) = \frac{2}{3} \frac{z}{(1-z)^2} - \frac{1}{3} \log (1-z) \qquad (2)$$

shows that this result is sh arp.

Theorem 3. (Distortion theorem).

Let $f \in K^*$. Then

(*i*) $[(3+r^2)] / [3(1+r)^3] \le | f'(z) | \le [3+r^2] / [3(1-r)^3]$

(*ii*) $\frac{2}{3} \frac{r}{(1+r)^2} + \frac{1}{3} \log (1+r) \le | f(z) | \le \frac{2}{3} \frac{r}{(1-r)^2} - \frac{1}{3} \log (1-r)$

The function f_0 given by (2) shows that these results are sharp.

By letting $r \longrightarrow 1^-$ in the lower bound of distortion result for f in theorem 3 (*ii*), we have the following covering result.

Theorem 4. Let $f \in K^*$. Then f maps the unit disc E onto a domain that contains the disc $|w| < 0.3977157$ and this result is sharp.

We shall need the following [10].

Lemma 1.

Let $g \in S^*$ in E and let $M(r, g) = \max_{z=r} [g(z)]$ and let α $(0 \leq \alpha \leq 1)$ be the order of $g(z)$. Then $\alpha = \lim_{r \to 1} (1-r)$ $[M'(r, g) / M(r, g)]$, where $M'(r, g)$ is the left derivative.

Theorem 5. (Arclength problem)

Let $f \in K^*$ and $g \in S^*$ in (1) be of order α $(0 \leq \alpha \leq 1)$. Let $L(r) = r \int_0^{2\pi} |f'(re^{i\theta})| \, d\theta$ denote the length of the image curve $C(r)$ of $|z| = r$. Then there is a $k = k(g)$ such that

$$L(r) \leq k + \frac{13}{\alpha} M(r, g)$$

Proof.

$$L(r) = \int_0^{2\pi} |zf'(z)| \, d\theta \leq \int_0^{2\pi} \int_0^{r} |(zf'(z))'| \, d\theta \, dr =$$

$$= \int_0^{2\pi} \int_0^{r} |g'(z) \, h(z)| \, d\theta \, dr,$$

where $g \in S^*$, $\operatorname{Re} h(z) > 0$ and $z = re^{i\theta}$.

Since $g \in S^*$; $\frac{z g'(z)}{g(z)} = H(z)$, $\operatorname{Re} H(z) > 0$. So

$$L(r) \leq \int_0^{r} \left(\int_0^{2\pi} |g(z) \, h(z) \, H(z)| \, d\theta \right) \frac{dr}{r}$$

$$\leq 2\pi \int_0^{2\pi} \frac{M(r, g)}{r} \left(\frac{1}{2\pi} \int_0^{2\pi} | h(z) |^2 d\theta \right)^{\frac{1}{2}}$$

$$\left(\frac{1}{2\pi} \int_0^{2\pi} | H(z) |^2 d\theta \right)^{\frac{1}{2}} dr ,$$

by using the Schwarz inequality.

Using a well-known result [10] for functions of positive real parts, we have

$$L(r) \leq 2\pi \int_0^r \left[\frac{M(r, g)}{r} \quad \frac{1 + 3r^2}{1 - r^2} \right| dr$$

$$\leq 2\pi \int_0^r \frac{2M(r, g)}{r(1-r)} dr \tag{3}$$

By Lemma 1, there is an $r_o(g)$ with $\frac{500}{501} < r_o < 1$ such that $(1 - \rho) \frac{M''(\rho, g)}{M(\rho, g)} > \frac{500}{501} \alpha$ for $r_o \leq \rho < 1$. Hence (3) shows that for $r > r_o$.

$$L(r) \leq 2\pi \int_0^{r_o} \frac{2M(\rho, g)}{\rho(1-\rho)} d\rho + 4\pi \frac{501}{500\, \alpha} \int_0^{r_o} M'(\rho, g)\, d\rho$$

$$\leq k + \frac{13}{\alpha} M(\rho, g)$$

In theorem 1, we have shown that $f \in K^*$ is close-to-convex. We now deal with the converse case as follows, see [6].

Theorem 6.

Let $g \in S^*$ and for $z \in E$, and let $\operatorname{Re} \frac{z f'(z)}{g(z)} > 0$. Then

$$\operatorname{Re} \frac{(z f'(z))'}{g'(z)} > 0 \text{ for } | z | < 2 - \sqrt{3}.$$

This result is sharp.

We now prove Livingstone's result [4] for the class K* .

Theorem 7.

Let $f \in K^*$ in E. Then $F(z) = \frac{1}{2}(z f(z))'$ is in K^* for $|z| < \frac{1}{3}$.

This result is sharp.

Proof. Let $f \in K^*$ with respect to a starlike function g. It is known [3] that $G(z) = \frac{1}{2}(zg(z))'$ is starlike in $|z| < \frac{1}{2}$. Now

$$\frac{(zf'(z))'}{g'(z)} = \frac{F'(z) + \frac{1}{z^2}\int_0^z F(z)\,dz - \frac{F(z)}{z}}{\frac{1}{z}G(z) - \frac{1}{z^2}\int_0^z G(z)\,dz} \tag{4}$$

Since $f \in K^*$, $\operatorname{Re} \frac{(zf'(z))'}{g'(z)} = \operatorname{Re} p(z) > 0, \; z \in E$. (5)

From (4) and (5), we have, by using integration by parts,

$$\frac{(zF'(z))'}{G'(z)} = p'(z)\left[\left(zG(z) - \int_0^z G(z)\,dz\right) / zG'(z)\right]$$

$$+ p(z) = p'[z]\left[\int_0^z zG'(z)\,dz\right] + p(z), \tag{6}$$

Now it is well-known [3] that $|p'(z)| \leq [2 \operatorname{Re} p(z)] / (1-r^2)$. Thus from (6), we have

$$\operatorname{Re} \frac{(zF'(z))'}{G'(z)} \geq \operatorname{Re} p(z)\left[1 - \frac{2}{1-r^2}\left|\frac{\int_0^z zG'(z)\,dz}{zG'(z)}\right|\right] \tag{7}$$

On using distortion theorems for the starlike function g, we have

$$\left| [(z.\ zG'(z)) / (\int_0^z zG'(z)\,dz)] \right| = \left| \frac{zg'(z))'}{g'(z)} + 1 \right| \geq \frac{2-4r}{1-r^2}$$

so that

$$\left| [\int_0^z zG'(z)\,dz] / (zG'(z)) \right| \leq \frac{r(1-r^2)}{2-4r} \tag{8}$$

From (7) and (8), we have

$$\text{Re}\ \frac{(zF'(z))'}{G'(z)} \geq \text{Re}\ p(z) \left[\frac{1-3r}{1-2r} \right] > 0 \text{ for } |z| < 1/3$$

The function f_0 given by (2) shows that the constant 1/3 cannot be improved.

We now consider the converse case of theorem 7 but in a generalized form as following.

Theorem 8.

Let $f \in K^*$ and for real $\alpha \neq 0$, let $F(z) = \frac{1}{\alpha} z^{1 - \frac{1}{\alpha}}$

$$\cdot \int_0^z z^{\frac{1}{\alpha} - 2} f(z)\,dz \text{ in E.}$$

Then $F \in K^*$ in E.

For $\alpha = \frac{1}{2}$, $F(z) = \frac{2}{z} \int_0^z f(z)\,dz$.

Proof.

Since $f \in K^*$, there exists a $g \in S^*$ such that $\text{Re}\ \frac{(zf''(z))'}{g'(z)} > 0$.

Also it is known [8] that $G(z) = \frac{1}{\alpha} z^{1-(1/\alpha)} \int_0^z z^{(1/\alpha)-2} g(z)\, dz$

belongs to S^*. Thus simple calculations yield

$$\frac{(zF'(z))'}{G'(z)}$$

$$= \frac{z^{\frac{1}{\alpha}} f'(z) + \left(\frac{1}{\alpha}-1\right)^2 \int_0^z z^{(1/\alpha)-2} f(z)\, dz - \left(\frac{1}{\alpha}-1\right) z^{(1/\alpha)-1} f(z)}{z^{(1/\alpha)-1} g(z) - \left(\frac{1}{\alpha}-1\right) \int_0^z z^{(1/\alpha)-2} g(z)\, dz}$$

$$= \frac{z^{\frac{1}{\alpha}} s''(z) - \left(\frac{1}{\alpha}-1\right) \int_0^z z^{(1/\alpha)-2} s''(z)\, dz}{\int_0^z z^{(1/\alpha)-1} t''(z)\, dz}$$, where $f(z)=s'(z)$ and $g(z)=t'(z)$.

Now

$$\frac{\left[z^{\frac{1}{\alpha}} s''(z) - \left(\frac{1}{\alpha}-1\right) \int_0^z z^{\frac{1}{\alpha}-1} s''(z)\, dz \right]'}{\left[\int_0^z z^{1-(1/\alpha)} t''(z)\, dz \right]'} = \frac{(z s''(z))'}{(t'(z))'}$$

$$= \frac{(zf'(z))'}{g'(z)}$$

Using the fact that $f \in K^*$ together with Libera's lemma [2], we have

$$\mathrm{Re}\ \frac{(zF'(z))'}{G'(z)} > 0, \quad z \in E,$$

and hence $F \in K^*$.

Remark 2. Theorem 7 has been generalized for $\alpha \in R$, see [11].

REFERENCES

1. W. Kaplan ; Clcse-tc-convex schlicht functions, Michigan Math. J. **1** (1952), 169-185.
2. R. J. Libera ; Scme classes of regular univalent functions, Proc. Amer. Math. Soc. **16** (1965), 755-758.
3. R. J. Libera ; Some radius of convexity problems, Duke Math. J. **31** (1964), 143-158.
4. A.E. Livingston ; On the radius of univalence of certain analytic functions, Proc. Amer. Math. Soc. **17** (1965), 352-357.
5. Z. Nehari ; Conformal mappings, McGraw-Hill, 1952.
6. K. Inayat Noor ; Some radius of convexity problems, C. R. Math. Rep. Acad. Sci. Canada, Vol. **4** (1982), 283-286.
7. K. Inayat Noor and F.M. Aloboudi ; A generalization of quasi-convex functions, Proc. Ist. Int. Conf. Math. Riyadh, 1982, to appear.
8. K. Inayat Noor, F.M. Aloboudi and N. Al-Dihan. On radius of univalence of convex combination of analytic functions, Exp. Math. (1984), to appear.
9. K. Inayat Noor and D. K. Thomas ; On quasi-convex univalent functions, Int. J. Math. and Math. Sci. **3** (1980), 255-266.
10. Ch. Rommerenke ; On starlike and close-to-eonvex functions, Proc. London Math. Soc. (3), **13** (1963), 290-304.
11. K. Inayat Noor, F.M. Aloboudi and N. Al-Dihan ; On the radius of univalence of convex combinations of analytic functions, Inter J. Math. and Math. Sci. (1983/84), to appear.

The Punjab University
Journal of Mathematics
Vol. XIV—XV (1981-82), pp. 193-198

L^p-CONDITION IN SPACES OF CONTINUOUS LINEAR MAPPINGS

By

M. NASIR CHAUDHARY
and
M. MAQBOOL ILAHI

Department of Mathematics
University of Engineering and Technology,
Lahore, Pakistan.

The duality of L^p-conditions was discussed by K.F. Ng in [5]. In this paper, we give necessary and sufficient conditions so that the space of continuous linear maps between two ordered Banach spaces will have the L^p-conditions.

Let X be an ordered Banach space with positive cone X_+. We say that X_+ is *generating* if, for each $x \in X$, there are $y, z \in X_+$ with $x = y - z$. X is *order complete* if every majorized subset of X has a least upper bound. X is *directed* (upwards) if, for each pair $x, y \in X$, there is $z \geq x, y$. The norm on X is monotone if $x, y \in X$ with $0 \leq y \leq x$ implies that $\| y \| \leq \| x \|$.

X is an *order unit normed* space if there is $e \in X_+$ such that for $x \in X$ there exists a positive integer n with $-ne \leq x \leq ne$, and the Minkowski functional of the order interval $[-e, e]$ defines tho norm on X.

We say that X is *regular* if it satisfies :

(R_1) : For each $x, y \in X$ such that $-x \leq y \leq x$, we have $\| y \| \leq \| x \|$; and

(R_2) : For each $x \in X$ and $\in > 0$, there is a $y \in X_+$ such that $y \geq x, -x$ and $\| y \| \leq \| x \| + \varepsilon$.

The following lemma is due to Bonsall [1] :

Lemma 1.

Let X be a real vector space with X_+ a wedge in X. Suppose P is a sublinear map from X into a complete vector lattice Y, and Q a superlinear map from X_+ into Y such that $Q(x) \leq P(x)$ for all $x \in X_+$. Then there is a linear map T from X into Y such that

$$T(x) \leq P(x), \quad x \in X.$$

and $Q(x) \leq T(x), \quad x \in X_+$.

Suppose $1 < p < \infty$. X satisfies

(L^p – *condition* (*i*) : if $x, y \in X_+$ then $\| x + y \|^p \geq \| x \|^p + \| y \|^p$.

L^p – *condition* (*ii*) : if $x, y \in X$ and $\varepsilon > 0$, then there exists $z \in X$ with $\| z \|^p \leq \| x \|^p + \| y \|^p + \varepsilon$ such that $z \geq x, y$.

It turns out that the two L^p –conditions are dual conditions [6 : 9.24, 9.25] :

Theorem 1. Suppose $\frac{1}{p} + \frac{1}{q} = 1$, then

(*a*) *X satisfies L^p—condition* (*ii*) *if and only if X^* satisfies L^q - condition* (*i*) :

(*b*) *X satisfies L^p - condition* (*i*) *if and only if X^* satisfies L^q - condition* (*ii*).

We extend these results to L (X, Y). First we give sufficient conditions so that L (X, Y) satisfies L^p – condition (*ii*) :

Proposition 1 :

Let Y be an order complete, order unit normed lattice. If X satisfies L^p – condition (*i*), *then L (X, Y) satisfies L^q – condition* (*ii*).

Proof.

Let F, G $\in$ L (X, Y). It suffices to show that there exists H $\in$ L (X, Y) with $\| H \|^q \leq \| F \|^q + \| G \|^q$ and $H \geq F, G$.

Let e be the order unit in Y. We define

$Q(z) = \text{Sup}\{F(x) + G(y) : x, y \in X_+, z = x + y\}, z \in X_+$

and $P(z) = \|z\| (\|F\|^q + \|G\|^q)^{1/q} . e, z \in X.$

Then Q is superlinear and P is sublinear. Moreover, if $z \in X_+$ and $z = x + y$, then

$$F(x) + G(y) \leq (\|F(x)\| + \|G(y)\|). e$$
$$\leq (\|F\| \|x\| + \|G\| \|y\|). e$$
$$\leq (\|F\|^q + \|G\|^q)^{1/q} (\|x\|^p + \|y\|^p)^{1/p} . e$$

(By Holder's inequality)

$$\leq (\|F\|^q + \|G\|^q)^{1/q} . \|z\| . e$$

(X satisfies L^p – condition (i))

$$= P(z).$$

i.e. $Q(z) \leq P(z), z \in X_+$. By Bonsall's lemma 1, there exists a linear map $H : X \longrightarrow Y$ such that

$$Q(z) \leq H(z), z \in X_+ \text{ and}$$
$$H(z) \leq P(z), z \in X.$$

From the definition of Q it is obvious that

$F(z), G(z) \leq H(z), z \in X_+$; *i.e.*

$F, G \leq H$. Moreover, for $w \in X$

$\|H(w)\| \leq (\|F\|^q + \|G\|^q)^{1/q} \|w\|$, since Y is 1 – normal.

This implies that $H \in L(X, Y)$; and

$$\|H\| \leq (\|F\|^q + \|G\|^q)^{1/q} ;$$

i.e. $\|H\|^q \leq (\|F\|^2 + \|G\|^q)$.

Thus $L(X, Y)$ satisfies L^q – condition (ii)

Now we consider sufficient conditions for $L(X, Y)$ to have L^p – condition (i).

Proposition 2.

Let the closed unit ball be directed in X. If Y satisfies (R_1) and L^p – condition (i), then $L(X, Y)$ satisfies L^p - condition (i).

Proof.

First we prove that if $F \in L(X, Y)_+$ and U is the closed unit ball in X, then $\| F \| = \operatorname{Sup} \{ \| F(x) \| : x \in U_+ \}$ where $U_+ = U \cap X_+$. For, if $x \in U$, then $-x \in U$ and there exists $z \in U_+$ with $x, -x \leq z$ *i.e.* $-z \leq x \leq z$, so that $-Fz \leq Fx \leq Fz$, and since Y satisfies (R_1) we obtain $\| Fx \| \leq \| Fz \|$.

Now let $S, T \in L(X, Y)_+$. We are required to show that

$$\| S + T \|^q \geq \| S \|^q + \| T \|^q .$$

If $S = 0$ or $T = 0$, then this inequality holds trivially. Therefore we can assume that $S \neq 0$, $T \neq 0$, so that we take a real number $\varepsilon > 0$ such that $0 < \varepsilon < \| S \|, \| T \|$. Then there are x, y in U_+ with

$$\| S \| - \varepsilon < \| S(x) \| \text{ and}$$

$$\| T \| - \varepsilon < \| T(y) \| .$$

Since U is directed there is $z \in U_+$ with $x, y \leq z$. Thus

$$\begin{aligned}(\| S \| - \varepsilon)^p + (\| T \| - \varepsilon)^p &< \| Sx \|^p + \| Ty \|^p \\ &\leq \| Sx + Ty \|^p \\ &\leq \| Sz + Tz \|^p \\ &\leq \| S + T \|^p \| z \|^p \\ &\leq \| S + T \|^p\end{aligned}$$

Passing to the limit as $\varepsilon \longrightarrow 0$, we obtain

$$\| S \|^p + \| T \|^p \leq \| S + T \|^p .$$

Next necessary conditions are considered for $L(X, Y)$ to have L^p – condition (i) or L^p – condition (ii) :

Proposition 3.

Let $L(X, Y)$ satisfy L^p – condition (i). Then X satisfies L^q – condition (ii) and Y satisfies L^p – condition (i).

Proof.

First we show that X^* satisfies L^p – condition (i).

Let $f, g \in X_+^*$ and $y \in Y_+$ with $\| y \| = 1$.

We define

$$F(x) = f(x).\, y, \quad x \in X \quad \text{and}$$

$$G(x) = g(x).\, y, \quad x \in X.$$

Then $F, G \in L(X, Y)_+$ and $\| F \| = \| f \|$, $\| G \| = \| g \|$. Thus

$$\begin{aligned} \| f \|^q + \| g \|^q &= \| F \|^q + \| G \|^q \\ &\leq \| F + G \|^q \\ &\leq \| f + g \|^q, \text{ since } \| F + G \| \leq \| f + g \|. \end{aligned}$$

Then X satisfies L^q – condition (ii) by Theorem 4.

To prove that Y satisfies L^p – condition (i), we take $a, b \in Y_+$. There exists $f \in X_+^*$ with $\| f \| = 1$. We define

$$A(x) = f(x) \,.\, a, \quad x \in X$$

$$B(x) = f(x) \,.\, b, \quad x \in X.$$

Then $A, B \in L(X, Y)_+$, $\| A \| = \| a \|$ and $\| B \| = \| b \|$.

$$\begin{aligned} \text{Thus } \| a \|^q + \| b \|^q &= \| A \|^q + \| B \|^q \\ &\leq \| A + B \|^q \\ &\leq \| a + b \|^q \end{aligned}$$

Proposition 4.

Let $L(X, Y)$ satisfy L^p – condition (ii). Then X satisfies L^q – condition (i) and Y satisfies L^p – condition (ii).

Proof.

Let $y \in Y_+$ with $\| y \| = 1$ and t be a positive continuous linear functional on Y with $\| t \| = 1$.

Let $f, g \in X^*$. We define

$$F(x) = f(x) \,.\, y, \quad x \in X$$

$$G(x) = g(x) \,.\, y, \quad x \in X.$$

Then F, G $\in$ L(X, Y) and $\|F\| = \|f\|$, $\|G\| = \|g\|$.

Since L(X, Y) satisfies L^p – condition (*ii*), so given $\varepsilon > 0$, there is H $\in$ L(X, Y) with H $\geq$ F, G and $\|H\|^p \leq \|F\|^p + \|G\|^p + \varepsilon$.

We define a functional h on X as $h(x) = t(H(x))$. Then h is linear, positive and $\|h\| \leq \|H\|$. Hence

$$\|h\|^p \leq \|H\|^p \leq \|F\|^p + \|G\|^p + \varepsilon$$
$$= \|f\|^p + \|g\|^p + \varepsilon$$

which implies that X* satisfies L^p – condition (*ii*). Therefore X satisfies L^q – condition (*i*).

Similarly, it can be easily seen that Y satisfies L^p – condition (*ii*).

REFERENCES

1. F.F. Bonsall. "The decomposition of continuous linear functionals into non-negative components". Proc. Univ. Durham Philos. Soc. A 13 (1957) 6-11.

2. E.B. Davies. "The structure and ideal theory of the predual of a Banach lattice", Tran. Amer. Math. Soc. **131** (1968) 544-555.

3. A. J. Ellis. "The duality of partially ordered normed linear spaces", J. Lond. Math. Soc. 39 (1964) 730-44.

4. G.J.O. Jameson. "Ordered linear spaces", Lecture Notes in Math. 104, Springer-Verlag, Berlin (1970).

5. K.F. Ng. "L^p – conditions in partially ordered Banach space", J. Lond. Math. Soc. (2) **5** (1972). 387-394.

6. Y.C. Wong and K.F. Ng. "Partially ordered Topological Vector Spaces", Clarendon Press, Oxford (1973).

CONTENTS